

RESULT	12
BG298479	
LOCUS	
DEFINITION	BC298479 900 bp mRNA EST 21-FEB-2001 602396747F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4511434 5', mRNA sequence.
ACCESSION	BC298479
VERSION	BC298479.1 GI:13063173
KEYWORDS	EST.
SOURCE	house mouse. Mus musculus
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
AUTHORS	1 (bases 1 to 900)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Plate: L1AM10394 row: n column: 11
High quality sequence stop: 676.

```

FEATURES             Location/Qualifiers
     source            1..900
                        /organism="Mus musculus"
                        /db_xref="taxon:10090"
                        /clone="IMAGE:4511434"
                        /clone_1bp="NIH_MGC_94"
                        /tissue_type="retina"
                        /lab_host="DH10B (phage-resistant)"
                        /note="Organ: eye; Vector: pCMW-SPORT6; Site_1: NotI;
                        Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
                        Average insert size 3.3 kb. Library enriched for
                        full-length clones and constructed by Life Technologies
                        Note: this is a NIH-MGC Library."
BASE COUNT            302 a      175 c      300 g      123 t
ORIGIN

```

[illegible]

RESULT	13
LOCUS	BF607177
DEFINITION	BF607177 830 bp mRNA
ACCESSION	U11000119 Mouse cDNA library ICRP522 Mus musculus
VERSION	BF607177.1 GI:13503669

KEYWORDS	EST.
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
	house mouse.
	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus 1 (bases 1 to 830)
	Yahyawi, M., Henmly, S., Neidhardt, L., Radelof, U., Hermann, B.G., Leberich, H. and O'Brien, J.
	Detection of a high number of novel genes in a 9-day mouse embryo cDNA library normalised by oligonucleotide fingerprinting
	Unpublished (2001)
	Contact: Henmly S

EST's are made from clones being representative of clone clusters.
Clone clusters were calculated from oligonucleotide fingerprints.
PCR primers
FORWARD: 5'-GAGCTATTCCAGAAGTGTGA-3',
BACKWARD: 5'-TAATACGACTCACTAATAGG-3',
Seq primer: 5'-ATTAGTGACACATATAC-3',
High quality sequence stop: 830.
Location/Qualifiers

```

source
1. .830
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="ICRFp522I0930"
/clone_id="Mouse 9-day fetus cDNA library ICRFp522"
/tissue_type="whole embryo"
/dev_stage="embryonic 9-day"
/lab_host="E.coli, XLI blue"
/note="Vector: pSVSPORT; Site_1: NotI; Site_2: SalI;
library preparation by oligo-dT priming of RNA. Clones can
be ordered from the Resource Center in Berlin,
http://www.rzpd.de."
268 a 161 c 271 g 125 t 5 others

```

	Query Match	63.3%	Score 140.6;	DB 11;	Length 830;
	Best Local Similarity	81.4%;	Pred. No. 2e-14;		
	Matches 193;	Conservative 0;	Mismatches 29;	Indels 15;	Gaps 2;
QY	1 gacatcagcgacagcgagaggaaagaccagctcgccgagtctcttcgcacgaatacagc	60			
Db	265 GACATTAAGTGACAGCGGAGAGAAACAGCTCAGTGTTCATCATCAGCAATTCAGCC	324			
QY	61 tccagttcttbaagaagaagagggaggaagag---gagggaggggaaaggagcaccc	117			
Db	325 TCAGGTTCTTAAGAGGGAGGAGGAAGAACAAGAGAAAGAAAGGAGGCCACC	384			
QY	118 agtgaagaalcagaggaaggaagaa-----gagggaggaaggaagacgcc	165			
Db	385 AGTGAAGAAATCATAGGAGGAAGAGAGAAAGADAGGAGGAGGAAGAAGAGAAACTGGG	444			
QY	166 agcaaacctctagagagatcatcagacgactcgcgcgaagaagttaattgaggagaatatg	222			
Db	445 AGCAACCTCTAGGAGGGCTCTTGAACACTCCGCACAAAGCAATGTATTGAGGAATG	501			

RESULT 14			
BF998861			
LOCUS	BF998861	456 bp	
DEFINITION	MR3-GN0152-101100-003-B03 GN0152	EST	24-JAN-2001
ACCESSION	BF998861	Homo sapiens	CDNA, mRNA sequence
VERSION	BF998861.1	GI:12434720	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

OY	120	tgaagaatca-gaggaagaaaggaaggaggggaggaagaaacgcggcgacactcttggg	178
Db	373	TGAATAATTCAGGAGGAGAAGAGAAAGAGAGAGAGACCGCAGCAACTCTGAGG	432
OY	179	aggatcacagaca	191
Db	433	AGGCATCAGAGCA	445
RESULT	8		
AM804794			
LOCUS	283 bp	mRNA	EST 16-MAY-2000
DEFINITION	OYA-UW0094-300300-156-g11 UW0094 Homo sapiens cDNA, mRNA sequence.		
ACCESSION	AM804794		
VERSION	AM804794.1	GI:7856768	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.		
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		
MEDLINE	20202663		
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-QV4-UW0094-300300-156-g11&t3=2000-03-30&t4=1) Seq primer: puc 18 forward High quality sequence start: 2 High quality sequence stop: 283. Location/Qualifiers		
FEATURES			
source	1..283	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone_lib="UM0094"	
		/dex_stge="Adult"	
		/note="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
BASE COUNT	87 a 49 c 114 g 33 t		
ORIGIN			
Query Match	75.3%; Score 167.2; DB 10; Length 283;		
Best Local Similarity	91.4%; Pred. NO. 8.7e-19;		
Matches 203; Conservative	0; Mismatches 8; Indels 11; Gaps 2;		
OY	10	gacacgaggaagaaacagcgtccgcagtcctctgtaagcagaatcagctcaggtct	69
Db	44	GTCACGAGAGAAAC--AGCTCGCCGACAGTCTCTGTCTCAGCGAATCAGGCTCAGATTCT	101
OY	70	gaggaagaagagagaggaaggaaggaaggaaggaaggaagcaccagtgaagaatca	129

[illegible]

Oy	61	tcaaggttcgcgaagaagaaggagagagggagagagagagagagagacaccat	120
Dd	305	TCAGGTCCTCGAAGAAGAAGACGACGAGAAAGAGACGACGACCACT	364
Oy	121	gaagaatacagag-----aggaagaagaagagagagagacccgacgaac	171
Dd	365	GAAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACCCGGACGAAC	424
Oy	172	tcttgagggagcatcagacagctctccgaagaagttaagtgaagaagaatg	222
Dd	425	TCTGAGAGGCATCAGACGACTCTCCGACAGAGTAGTGTAGGCCAGAAATG	475
RESULT	6		
LOCUS	B1018825/c		
ACCESSION	B1018825	493 bp	mRNA
VERSION	B1018825.1	GI:14425455	EST
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.		
	1 (bases 1 to 493)		
	Dias Neto, E., Garcia Correia, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. P., Goldmann, G. H., Carvalho, A. F., Matsukuma, A., Bais, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.		
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	97 (??),	3491-3496 (2000)
MEDLINE	20202663		
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?IL=113&t2=IL3-MT0267-050101-435-H02&t3=2001-01-05&t4=1) Seq primer: puc 18 forward High quality sequence start: 13 High quality sequence stop: 283.		
FEATURES			
Source	Location/Qualifiers		
	1..493	/organism="Homo sapiens"	
	/db_xref="taxon:9606"		
	/clone_lib="MT0267"		
	/dev_stage="Adult"		
	/note="Organ: marrow; Vector: puc18; Site_1: Smal; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 1967,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."		
BASE COUNT	66 a	178 c	90 g
ORIGIN			159 t
Query Match	81.4%; Score 180.6; DB 11; Length 493;		
Best Local Similarity	90.0%; Pred. NO. 5.3e-21;		
Matches	208; Conservative	0; Mismatches	14; Indels
		9; Gaps	1;
Oy	1	gacatcagcgacagcgagagaaacacgctcgccgagctctgcgcagcagatcagc	60

Db	347	GGCATCATTTGACAGCGAGGAAGACCAGCTGGCCGCAAGTCTCCTCAATTGTATTCAGAGT	288
OY	61	tcaagttcttgaggaaagaaggaaggaaggaaggaaggaaggaaggaaggaaggaagcaccagt	120
Db	287	CCAGGTTCTTGAGGAATAACAGGACGAGGAAAGAGAGAGAGAGAGAACGCCGACACACT	228
OY	121	gaagatcac-----gaggaggaagaaggaaggaaggaaggaaggaaggaagccgcacac	171
Db	227	GAAAGAATCAGAGAGAGAGAGAGAGATTTGGAAAGAGAGAGAGAGAGACCGCGCACAC	168
OY	172	tctggaggaagcatcgagcagtcgtccgaagaagtaagttaggaagaagaaatg	222
Db	167	TCTGAGGAGGACATCAGAGCAGCTCTCCGCAAGTAAGTAGAGAGAAAGAAATG	117
RESULT	7		
BG822824		BG822824	662 bp mRNA EST 22-MAY-2001
LOCUS		602727816P1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4867443 5,	
DEFINITION		mRNA sequence.	
ACCESSION		BG822824	
VERSION		BG822824.1 GI:14170411	
KEYWORDS		EST.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
TITLE		NIH-MGC http://mgc.nci.nih.gov/.	
JOURNAL		1 (bases 1 to 662)	
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: rsraus@nhi.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: NIH Intramural Sequencing Center Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LNCML735 row: p column: 04 High quality sequence stop: 601. Location/Qualifiers 1..662 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4867443" /clone_lib="NIH MGC 15" /tissue_type="adenocarcinoma cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)".	
FEATURES			
SOURCE			
Query Match		76.1%; Score 169; DB 11; Length 662;	
Best Local Similarity		99.0%; Fred. No. 4.2e-19;	
Matches 191; Conservative		0; Mismatches 0; Indels 2; Gaps 2;	
OY	1	gacatcagcgaagaaggaaggaagcagctcgccgagctcctctgacagaaatcagcgc	60
Db	253	GACATCAGCGAGGAGGAGAGGAAGACCAAGCTGGCGCAGTCTCTCAGACAATCAGGC	312
OY	61	tcaagttcttgaggaaagaaggaaggaaggaaggaaggaaggaaggaaggaagcacc-ag	119
Db	313	TCAGGTTCTTGAGGAAGAGAGAGAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGC	372

Dd	276	GAAATCAGCGACGAGGAGGAAGAACAACGCTCGGGCCGAGTCTTGTCATCGAGGAATTCAGGC	335
Oy	61	tcaaggtcttcaggaaagaagaggaaggaagaagaagaagaagaagccaccagt	120
Dd	336	TCAGGTCTTCGAGGAAGAAGAGAGAGAGGAAGAAGAGGAGGAGGAAGAAGGAGCACCACT	395
Oy	121	gaagaatca-----gagaggaagaagaagaagagagagagagacccgcagacaac	171
Dd	396	GAAGAATCATGAGAGGAGGAGGAGGAGGAGGAAGAAGAGGAGGAGGAGGAGACCCTGCACAC	455
Oy	172	tcttgaggaggcatcagagcacctctgccgaaaagttaagtggagaagaaty	222
Dd	456	TCTGAGGAGGCAATCAGACAGCTCTGCCAAGAACAGTAAGTAGGAGGAAGAAATG	506
RESULT	4		
LOCUS	B1261348	345 bp	mRNA
DEFINITION	602953350P1 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5087556 5'	EST	17-JUL-2001
ACCESSION	B1261348		
VERSION	B1261348.1	GI:14820527	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgs.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		

```

FEATURES
SOURCE
118 a 57 c 131 g 39 t
Tissue Procurement: Dr. Louis Staudt, M.D., Ph.D. CDNA Library
Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LUCM1845 row: c column: 13
High quality sequence stop: 345.
Location/Qualifiers
1. 345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5087556"
/clone_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pOT8; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(c). Size selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

	Query Match	88.8%	Score 197.2;	DB 11;	Length 345;
	Best Local Similarity	94.4%;	Pred. NO. 1e-23;		
	Matches 219; Conservative	0;	Mismatches 3;	Indels 10;	Gaps 1;
Oy	1 gacatcagcgaaagagaagaaaccagctcygscgaagtccctcgtcagacgaalcaagc	60			
	13 GACATCAGCGAAGGAGGAGAAGCAACGCTGGCCGAGTCTCTCTCAGCGGAATCAGGC	72			
Db	61 tcaagttctgaggaagaaaggagagagaaagagagagagagaaaggagcaccaat	120			

[illegible]

RESULT	5
BF156008	
LOCUS	BP156008 516 bp mRNA EST 30-OCT-2000
DEFINITION	RCO-HR09G95-2709000-033-d08 HR0955 Homo sapiens CDNA, mRNA sequence.
ACCESSION	BF156008
VERSION	BF156008.1 GI:11051191
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 516)
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R. Negal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsushima,A., Bala,G.S., Simpson,D.H. Brunshtein,A.P., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
COMMENT	Contact: Simpson A.J.G.

```

FEATURES
source
    Ludwig Institute for Cancer Research
    Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
    Brazil
    Tel: +55-11-2704922
    Fax: +55-11-2707001
    Email: asimpson@ludwig.org.br
    This sequence was derived from the FAPESP/LICR Human Cancer Genome
    Project. This entry can be seen in the following URL
    (http://www.ludwig.org.br/scripts/gethtml2.pl?fl=at2-RC0-HR0955-2707001-033-d08&f3=2000-09-27&f4=1)
    Seq primer: puc 18 forward
    High quality sequence start: 56
    High quality sequence stop: 516.
    Location/Qualifiers
        1..516
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_id="HR0955"
            /dev_stage="Adult"
            note="Organ: head/neck; Vector: puc18; Site.1: Sma1;
            Site.2: Sma1; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            into the pUC 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."
BASE COUNT
155 a 109 c 175 g 77 t
ORIGIN

```

Query Match	88.6%	Score 196.6	DB 11	Length 516
Best Local Similarity	94.4%	Pred. No. 1.2e-23		
Matches 218	Conservative	0	Mismatches 4	Indels 9
				Gaps 1
Qy	1	gacatcagcgcagacgcgagaggaagacacagctctgcgcgagctctctcgttcagcagaalcaagc	60	
Db	245	GACATCAGCGCAGCGAGGAGGAAGACCACTCTGGCCGCGATCTCTGTCTCAGCGGAATTCAGGC	304	

```

/issue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: Not1; Site_2: Sal1"
BASE COUNT      230 a      155 c      254 g      100 t      1 others
ORIGIN

```

Query Match	90.7%;	Score 201.4;	DB 10;	Length 740;
Best Local Similarity	95.7%;	Pred. No. 2e-24;		
Matches 221; Conservative	0;	Mismatches 1;	Indels 9;	Gaps 1;

1 gacatcagcgcagcgcgagaggaaccagctcggccgagtcctcgtcagcagaatcagcgc 60
 |||||
 44 GACATCAGCGCAGCGGAGAGGACCACTCGGCGGAGTCTCGTCAAGCGAATCAAGCG 100
 |||||

[illegible]

121 gaaagatca-----gagagaggaagagagagagagagacccgcagcaac 17

172 tctgaggaagcattcagagcactctgccgaagaaglaagttaggaagaatb 222

224 TCTGAGGAGCATCAGAGCAGTCTGCCGAAGAGTAACTGACGGAAGAAATG 274

RESULT 2
BE740559

LOCUS	BE440559	1063 bp	MRNA	EST	15-SEP-2000
DEFINITION	601595656f1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3949653 5 mRNA sequence.				

```

VERSION      BE740559.1  GI:10154551
KEYWORDS     EST.
SOURCE       human.

```

ORGANISM	Homologous
<i>Homo sapiens</i>	100
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	99
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo	98
100	97

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

CONTACT: ROBERT STRAUSBERG, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/OTIP

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found at <http://www.ncbi.nlm.nih.gov/Genbank/MGC/>

Found through the I.M.A.G.E. Consortium/ENL at: [Image.Lnl.g](#)
Plate: LLCM814 row: p column: 14
High quality sequence stop: 741.

```
source      1. 1063
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
```

```

/clone_image="3945093"
/clone_id="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"

```

/note="Organ: ovary; Vector: pOMB7; Site 1: XhoI; SI
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'

insert size 1.8kb. Library constructed by Ling Hong the laboratory of Gerald M. Rubin (University of California, Berkeley) and sent to me for sequencing.

BASE COUNT	ORIGIN	(Stratagene) and Superscript II RT (Life Technologies)
351 a	203 c	402 g 107 t

Query Match	90.7%;	Score 201.4;	DB 10;	Length 1063;
Best Local Similarity	95.7%;	Pred. No. 1,9e-24;		
Matches 221;	Conservative	0;	Mismatches 1;	Indels 9;
				Gaps 1

Qy 1 gacatcagcgacagcgagaggaagaccagctcgccgagccctcgtcagcagaatcaagc 60
 408 GACATCAGCGACAGCGAGAGGAGAGACCAGCTCGCCGAGTCTTCGTACAGCGAATCAGGC 467

[illegible]

QY 121 gaagaatca-----gaggaaggaaggaaaggagggaggaaggaaccgcaghaac 171

QY 172 tctgagagcatcagagcagctctgcgagaagtcagtgaggaagaatg 222

Db 588 TCTGAGCAGGCATCAGCAGCAGTCTGCCGAGGAGTAAAGTGAGGAGAGAAATG 638

DEFINITION	LOCUS	1340 bp	mrna	EST	30-JAN-2001
6022727879a1	6022727879a1				

	mRNA sequence.
ACCESSION	BG107702
VERSION	BG107702.1 GI:12601548

SOURCE ORGANISM	human, Homo sapiens
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostei	

REFERENCE
AUTHORS
1 (bases 1 to 1340)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
NIH-MGC <http://mgc.ncl.nih.gov/>.

JOURNAL COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: carabs-rr@mail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

FEATURES
Plate: LLAM10014 Row: 1 Column: 19
High quality sequence stop: 667.
Location/Qualifiers

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:4365474"
```

```

/clone_id="NH MCC_86"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"

```

Site_2: Salt; Cloned unidirectionally; oligo-dT prime
Average insert size 1.533 kb. Library enriched for
full-length clones and constructed by Life Technologies

BASE COUNT	NOTE: THIS IS A NIH-MGC LIBRARY.			
ORIGIN	467 a	272 c	450 g	151 t

Query Match	90.78;	Score 201.4;	DB 11;	Length 1340;
Best Local Similarity	95.78;	Pred. No. 1.9e-24;		

QY	1	gacatcagcgacagcgagagaacacagctcggccgaatctctcgtcagcagaaatcagc	60
matches	221	CONSERVATIVE	0
mismatches	1	INDELS	9
gaps			

OY	31	tcggccgagcttcctcgtcagcagaatacgaagtccaagttcttgaaagaaagagaggag	90
Db	1191	TCGGCCAAATCACCGGTGATNACTACCACTGCCAGAAAGAAGAAGAAGAGAGAGAG	1132
OY	91	gaagaaggagagagagagagagaccacagttaaatactcagagagagaaagagagagag	150
Db	1131	GAGGAAGAAGAAAGCAGTAGAAGAAGTGGAAAGTAGAAGTAGATGTCGAATGTG	1072
OY	151	gagagagagacccgcagcaactcttgtagggatctaagcagctcgcgccgaagaagt	206
Db	1071	GAGAACTACTAGATTTTAGTGCCAGAAAGAACCTAGATGTTGGTTGCCGAAGAAGT	1016

```

RESULT 11
US-08-396-001-13/c
; Sequence 13, Application US/08396001
; Patent No. 5919618
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Nicnor
; APPLICANT: Claus, James
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OF INVENTION: Genes Determining Cellular Senescence in
; TITLE OF INVENTION: Yeast
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/396,001
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MTI-6408A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2150 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 563..1987
; IS-08-396-001-13

```

Query Match	25.98;	Score 57.6;	DB 2;	Length 2150;
Best Local Similarity	58.0%;	Pred. No. 2e-05;		
Matches 102;	Conservative	0;	Mismatches 74;	Indels 0;
				Gaps 0;

QY 31 tggcgcgagjctccgcgcacgacgaatcaggtcttgaggaagaagagagag 90
 Db 1191 TCGCCAAATCACCTGATGATCTACTACCACTGCCAGAAAGAAGAGAGAGAGAGAG 1132
 QY 91 gaagagagagagagagagagagagcaccagtgagaatcagagagagagagagag 150

Accession	Sequence	Length
Db 1131	GAGGACAGCGAGAGAGAGTAGAAGAAAGTCGAAAGTAGAAGTAGAGAGTAGAGTCGAAAGTG	1072
Qy 151	gagagagagacccgcagcaactctgaggagcactcagccagctctgcgaagaagt	206
Db 1071	GAGGAGCTACTAGTTTAGTGGCGAAGAAAGCCCTTGATTTGGGTTGGCGGAACAAGT	1016

```

RESULT 12
US-09-323-433A-13/c
; Sequence 13, Application US/09323433A
Patent No. 6218512
GENERAL INFORMATION:
APPLICANT: Gaudente, Leonard P.
APPLICANT: Austriaco Jr., Nicanor
APPLICANT: Claus, James J.
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESENCE IN
TITLE OF INVENTION: YEAST
FILE REFERENCE: 0050.1491-003
CURRENT APPLICATION NUMBER: US/09/323,433A
PRIORITY FILING DATE: 1999-06-01
PRIORITY FILING DATE: 1995-02-28
PRIORITY APPLICATION NUMBER: PCT/US94/09351
PRIORITY FILING DATE: 1994-08-15
PRIORITY APPLICATION NUMBER: US 08/107,408
PRIORITY FILING DATE: 1993-08-16
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13
LENGTH: 2150
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
FEATURE:
NAME/KEY: CDS
LOCATION: (363)...(1987)
OTHER INFORMATION: SAG1
US-09-323-433A-13

```

[illegible]

```

RESULT 13
US-08-728-323A-1
? Sequence 1, Application US/08728323A
? Patent No. 594676
?
? GENERAL INFORMATION:
?
? APPLICANT: Chang, Yuan
?
? APPLICANT: Bohenzky, Roy A.
?
? APPLICANT: Russo, James J.
?
? APPLICANT: Edelman, Isidore S.
?
? APPLICANT: Moore, Patrick S.
?
? TITLE OF INVENTION: Immediate Early Protein From Kaposi's
?
? TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
?
? TITLE OF INVENTION: Encoding Same And Uses Thereof
?
? NUMBER OF SEQUENCES: 21
?
? CORRESPONDENCE ADDRESSES:

```


DY 70 ggaggaagaagagggggaggaaggaaggaaggaaggaagcaaccagtgaagatca 129
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 GAGGAGGAGCAGCAGGAGGAGGAGGAAGAGCAGCAGCAGAAGAAGAGGACGACAG 167

[illegible]

RESULT 4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 16:24:14 ; Search time 152.61 Seconds
(without alignments)
329.455 Million cell updates/sec

Title: US-09-915-060-1

Perfect score: 222
Sequence: 1 gacatcagcagcagcagag.....aagtaagtgcagaagaatg 222

Scoring table: IDENTITY-MDC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCRTUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67.4	30.4	7218	1	US-08-332-463-14 Sequence 14, Appl
2	62.6	28.2	51259	3	US-08-781-891-209 Sequence 209, App
3	60.6	27.3	3211	2	US-08-574-959A-8 Sequence 8, Appl
4	60.6	27.3	3211	4	US-09-357-014-8 Sequence 8, Appl
5	60.6	27.3	3901	2	US-08-574-959A-6 Sequence 6, Appl
6	60.6	27.3	3901	4	US-09-357-014-6 Sequence 6, Appl
7	58	26.1	1931	2	US-09-130-114-2 Sequence 2, Appl
8	57.8	26.0	15378	3	US-08-785-420-1 Sequence 1, Appl
9	57.8	26.0	43795	3	US-08-742-185-101 Sequence 101, App
10	57.6	25.9	2150	2	US-08-861-464-13 Sequence 13, Appl
11	57.6	25.9	2150	4	US-08-396-001-13 Sequence 13, Appl
12	57.2	25.8	3489	2	US-08-728-323A-1 Sequence 1, Appl
13	57.2	25.8	3489	4	US-08-728-323A-1 Sequence 2, Appl
14	57.2	25.8	32207	2	US-08-770-379-20 Sequence 20, Appl
15	57.2	25.8	32207	4	US-08-757-669A-20 Sequence 4, Appl
16	56.4	25.4	2340	3	US-09-022-983-4 Sequence 3, Appl
17	56.4	25.4	2477	4	US-09-490-692-3 Sequence 4, Appl
18	56.2	25.3	9636	1	US-08-323-170B-1 Sequence 1, Appl
19	55.4	25.0	16442	3	US-08-781-891-208 Sequence 208, App
20	53.8	24.2	2301	5	US-08-506-691B-23 Sequence 23, Appl
21	53.8	24.2	2301	5	US-08-506-691B-23 Sequence 78, Appl
22	53.6	24.1	19056	4	US-09-272-032-8 Sequence 8, Appl
23	53	23.9	1276	4	US-09-177-325-2 Sequence 2, Appl
24	53	23.9	1276	4	US-09-411-812A-2 Sequence 2, Appl
25	53	23.9	1276	4	US-09-590-113-2 Sequence 1, Appl
26	52.4	23.6	1052	1	US-08-466-603-1 Sequence 1, Appl
27	52.4	23.6	1052	1	US-08-314-503A-1 Sequence 1, Appl

28	52.4	23.6	1052	1	US-08-468-066-1 Sequence 1, Appl
29	52.4	23.6	1052	2	US-08-466-717-1 Sequence 1, Appl
30	52.4	23.6	1052	3	US-08-466-743-1 Sequence 1, Appl
31	52.4	23.6	1052	5	PCT-US95-12414-1 Sequence 1, Appl
32	51.8	23.3	759	1	US-08-466-603-4 Sequence 4, Appl
33	51.8	23.3	759	1	US-08-314-503A-4 Sequence 4, Appl
34	51.8	23.3	759	1	US-08-468-066-4 Sequence 4, Appl
35	51.8	23.3	759	2	US-08-468-066-4 Sequence 4, Appl
36	51.8	23.3	759	3	US-08-466-717-4 Sequence 4, Appl
37	51.8	23.3	759	5	PCT-US95-12414-4 Sequence 4, Appl
38	51.8	23.3	960	1	US-08-466-603-3 Sequence 3, Appl
39	51.8	23.3	960	1	US-08-314-503A-3 Sequence 3, Appl
40	51.8	23.3	960	1	US-08-468-066-3 Sequence 3, Appl
41	51.8	23.3	960	2	US-08-466-717-3 Sequence 3, Appl
42	51.8	23.3	960	2	US-08-466-743-3 Sequence 3, Appl
43	51.8	23.3	960	5	PCT-US95-12414-3 Sequence 3, Appl
44	51.2	23.1	5183	1	US-08-459-568-3 Sequence 3, Appl
45	51.2	23.1	5183	2	US-08-399-411-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

AC	AAQ39212;
XX	
DT	21-JUL-1993 (first entry)
XX	
XX	CENP-B CDNA.
XX	
KW	Centromere; autoantigen; anticentromere; ACA; CREST; chromosomal;
KW	auto immune disease; calcinosis; Raynaud's phenomenon; esophageal;
KW	disability; scleroactyl; elanglectasiae; scleroderm; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1785
FT	/*tag= a
FT	39..483
FT	/*tag= b
FT	/note= "from clone CNP-B3"
FT	484..1334
FT	/*tag= c
FT	/note= "from clone CNP-B2"
FT	1335..2643
FT	/*tag= d
FT	/note= "from clone CNP-B1"
XX	
PN	US5196307-A.
XX	
PD	23-MAR-1993.
XX	
XX	29-MAR-1988; 88US-0174854.
XX	
XX	29-MAR-1988; 88US-0174854.
PR	
XX	(UYJO) UNIV JOHNS HOPKINS.
PA	
XX	
PI	Cleveland D, Earnshaw W, Sullivan KF;
XX	
DR	WPI: 1993-116791/14.
XX	
DR	P-PSDB; MAR34936.
XX	
XX	
PT	Human centromere antigen CENP-B and its cDNA - for detecting
PT	auto-antibodies to CENP-B in auto-immune disease diagnosis
XX	
PS	Claim 1; Fig 1; 20pp; English.
XX	
CC	Screening of a human endothelial cell cDNA library in lambda g11
CC	with serum G6 provided a positive clone, lambda CENP-B1, contg. a
CC	cDNA insert of 1,282 bp. Phage clones contg. longer cDNA inserts
CC	were obtd. by screening a cDNA library constructed from human
CC	placental mRNA, clones lambda CENP-2.4. The clones were found to
CC	have overlapping sequences and together made up the complete CENP-B
CC	cDNA sequence. The cDNA encodes a centromere polypeptide comprising
CC	an epitope which is reactive with samples contg. anti-centromere
CC	autoantibodies (ACAs) and is destroyed by disruption at a site 42
CC	amino acids from the C-terminus of CENP-B or immunologically binds
CC	the monoclonal antibody produced by hybridoma cell line ATCC HB 9667.
CC	The cDNA may be used to transform cells which may be used to produce
CC	recombinant CENP-B which may be used as an assay reagent for detection
CC	of autoantibodies in the diagnosis of autoimmune diseases, esp.
CC	calcinosis/Raynaud's phenomenon/esophageal dysmobility/scleroactyl/
CC	elanglectasiae (CREST) variant of scleroderma.
XX	
XX	
Sequence	2643 BP; 512 A; 801 C; 878 G; 451 T; 1 other;

	Query Match	29.1%	Score 64.6;	DB 14;	Length 2643;
	Best Local Similarity	64.2%;	Pred No.	3.5e-06;	
	Matches	97;	Conservative	0;	Mismatches 54; Indels 0; Gaps 0;
Oy	70	gaggaagaaggaggaagaagaagaagaagaagaagacaccagtagaatca	129		
Db	1201	gaggaagaaggaggaagaagaagaagaagaagaagaagggttgaagaagaagaag	1260		

```

Oy    130 gagggaggaggaagaagaaagaacgacgcgcgaactctgaggatcatcagag 189
      ||||| ||||| ||| | | | | | | | | | | | | | | | | | | | | | | |
Db    1261 gaggaggagagggaggaggaggaggaagaaggggggaagagaattggggaggaaa 1320
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy    190 cagtctgccgaagaaqtaagtcgaagaagaaa 220
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    1321 gagagggtcgaggaggaggagtgatgttggata 1351
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT  14
ID       AAA64660 standard; cDNA; 3717 BP.
XX
AC       AAA64660;
XX
DT       02-JAN-2001 (first entry)
XX
DE       DNA encoding centromere autoantigen B (CENP-B) protein.
XX
KW        Melanoma; thyroid tumor; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KM        rectal cancer; lung cancer; breast cancer; colon cancer;
KN        centromere autoantigen B; CENP-B protein; ss.
OS        Homo sapiens.
XX
PN        WO200050595-A2.
PD        31-Aug-2000.
XX
PF        25-FEB-2000; 2000MO-US04929.
XX
PR        25-FEB-1999; 99US-0257417.
XX
PA        (GOUT/) GOUT I.
PA        (RODN/) RODNIN N.
PA        (FILO/) FILOMENKO V.
PA        (MATS/) MATSUKA G.
PA        (SCAN/) SCANLAN M.
PA        (OLDL/) OLD L.
PA        (BILY/) BILYSKY B.
XX
PI        Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;
PI        Bilysky B;
DR        WPI: 2000-572092/53.
XX
PT        Novel isolated nucleic acid molecules for diagnosing and treating -
PT        melanoma, thyroid tumors, rectal, lung, breast and colon cancers -
PS        Claim 17; Page 86-87; 94pp; English.
XX
CC        The specification describes polynucleotides which are associated with
CC        melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,
CC        Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for
CC        diagnosing and treating a patient with melanoma, thyroid tumour,
CC        rectal cancer, lung cancer, breast cancer or colon cancer. The
CC        present sequence represents a polynucleotide of the invention.
XX
Sequence 3717 BP; 705 A; 1133 C; 1214 G; 665 T; 0 other.
```

[illegible]

[illegible]

FT		/product= rianodin receptor
FT	GC_signal	351..356
FT		/tag= b
FT		/note= "feature unlabelled in specification"
FT	GC_signal	407..452
FT		/tag= c
FT		/note= "feature unlabelled in specification"
FT	CANAL_signal	392..396
FT		/tag= d
FT		/note= "feature unlabelled in specification"
FT	misc-feature	503..510
FT		/tag= e
FT		/note= "feature unlabelled in specification"
FT	misc-feature	514..521
FT		/tag= f
FT		/note= "feature unlabelled in specification"
FT	polyA_signal	15388..15393
FT		**tag= g
XX		
XX	JP03011098-A.	
XX		
PD	18-JAN-1991.	
XX		
PF	07-JUN-1989;	89JP-0144569.
XX		
PR	07-JUN-1989;	89JP-0144569.
XX		
PA	(MITU) MITSUBISHI KASEI CORP.	
DR	WPI. 1991-062003/09.	
XX	P-PBDB; AAR10834.	
PT	New rianodin receptor, genes encoding it and its prepn. - useful as	
PT	calcium release modulator for tranqullisers and for assaying	
PS	calcium antagonists.	
XX		
CC	Disclosure; Fig 1; 18pp: Japanese.	
XX		
CC	RNA contg. poly(A) was prepared from rabbit skeletal muscle	
CC	endoplasmic reticulum. From the obtd. poly(A) mRNA, a cDNA bank	
CC	corresp. to it was prepared by random primer method, oligo (dT)	
CC	primer method, and primer extension method. A cDNA was obtained by	
CC	screening with a DNA probe (see AAQ10614-15). By introducing the obtd.	
CC	cDNA into an expression vector, vector PRS7 was formed.	
CC	The product is said to be involved in calcium release from	
CC	sarcoplasmic reticulum which triggers constriction of skeletal muscle.	
CC	Therefore, the receptor is useful as tranqulliser and assay series for	
CC	screening of calcium antagonist.	
XX		
SQ	Sequence 15672 BP; 3139 A; 4909 C; 4955 G; 2669 T; 0 other;	
	Query Match	29.3%; Score 65; DB 12; Length 15672;
	Best Local Similarity	65.5%; Pred. No. 3.6e-06;
	Matches 95; Conservative	0; Mismatches 50; Indels 0; Gaps 0;
QY	70 gaagaaagaaggaaggaggaagaagaaggaaggaagcaccgtgaagaatca	129
Db	6221 gaagaaagaaggaaggaggaagaagaaggaaggaaggaacgaagaagaagaagaa	6280
QY	130 gaagaaagaaggaagaaggaaggaaggaagccgcagacaactctgaagagcatcagag	189
Db	6281 gaagacgaggaaggaaggaaggaaggaagtgtcgtaaaagaaggaaggaagccccgaa	6340
QY	190 cagtcgccgaagaagtaagtgg 214	
Db	6341 ggaggaaaaagaagattcgaggagg 6365	


```
FT /note= "CDS does not include start codon"
FT /partial
FT mutation (931..932, AGAG)
FT /tag= b
FT /note= "This deletion results in the change of
FT amino acids and ends in a premature stop codon"
XX
PN WO200138578-A1.
XX
XX 31-MAY-2001.
XX
XX 21-NOV-2000; 2000WO-US31940.
XX
XX 24-NOV-1999; 99US-0167365.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
XX
XX WPI: 2001-367707/38.
XX
XX P-PSDB; AAE02399.
XX
XX Identifying dogs with or carrying X-linked progressive retinal atrophy
XX by detecting retinitis pigmentosa GTPase regulator gene mutation,
XX useful when breeding Husky, Samoyed and Miniature Schnauzer -
XX
XX Claim 75; Page 34-35; 88pp; English.
XX
XX The invention relates to a method for identifying dogs which are
XX genetically normal, are carriers of, or are affected with X-linked
XX progressive retinal atrophy (XLPRA), by testing a biological sample with
XX genetic markers that co-segregate with a XLPRA gene locus. The invention
XX also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
XX mutants and their corresponding nucleic acid molecules. The mutated RPGR
XX genes are responsible for the XLPRA in dogs. Methods are used to select
XX dogs for breeding so that dogs carrying the mutated locus are eliminated
XX from the breeding stock. The method particularly applies to Siberian
XX Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
XX the disease is X-linked. XLPRA1 type is identified in Siberian Huskies,
XX and Samoyeds, while XLPRA2 type is identified in Miniature Schnauzers.
XX The present cDNA sequence is the mutant open reading frame (ORF) 15
XX encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant
XX found in XLPRA2-affected dogs. This mutant is obtained by deleting
XX 'GA' nucleotide bases from position 932 to 933 of the wild-type
XX canine RPGR cDNA.
XX
XX Sequence 2803 BP; 1028 A; 287 C; 939 G; 549 T; 0 other:
SQ
Query Match 30.6%; Score 68; DB 22; Length 2803;
Best Local Similarity 62.2%; Pred. No. 5.9e-07;
Matches 107; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 48 agcagatcaggtcaggtctctgaggaagaagaggaaggaaggaaggaagga 107
DB 1069 agggagaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 1128
QY 108 agggagaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag 167
DB 1129 agggagaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 1188
QY 168 caactcgaagcagcagcagcagcagcagcagcagcagcagcagcagcag 219
DB 1189 ggaagggagcaaggaaggaaggaaggaaggaaggaaggaaggaagga 1240
RESULT 8
AAD07130
ID AAD07130 standard; cDNA; 2805 BP.
XX
XX AAD07130;
XX
XX 06-AUG-2001 (first entry)
```

```
XX
XX Canine retinitis pigmentosa GTPase regulator (RPGR) cDNA.
DE
XX Dog: X-linked progressive retinal atrophy; XLPRA; genetic marker.
XX retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
XX Miniature Schnauzer; ss.
XX
XX Canis familiaris.
XX
XX Key Location/Qualifiers
FH 1..1569
FT CDS
FT /tag= a
FT /product= "Canine retinitis pigmentosa GTPase regulator"
FT /note= "CDS does not include start codon"
XX
XX WO200138578-A1.
XX
XX 31-MAY-2001.
XX
XX 21-NOV-2000; 2000WO-US31940.
XX
XX 24-NOV-1999; 99US-0167365.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
XX
XX WPI: 2001-367707/38.
XX
XX P-PSDB; AAE02399.
XX
XX Identifying dogs with or carrying X-linked progressive retinal atrophy
XX by detecting retinitis pigmentosa GTPase regulator gene mutation,
XX useful when breeding Husky, Samoyed and Miniature Schnauzer -
XX
XX Claim 71; Page 30-31; 88pp; English.
XX
XX The invention relates to a method for identifying dogs which are
XX genetically normal, are carriers of, or are affected with X-linked
XX progressive retinal atrophy (XLPRA), by testing a biological sample with
XX genetic markers that co-segregate with a XLPRA gene locus. The invention
XX also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
XX mutants and their corresponding nucleic acid molecules. The mutated RPGR
XX genes are responsible for the XLPRA in dogs. Methods are used to select
XX dogs for breeding so that dogs carrying the mutated locus are eliminated
XX from the breeding stock. The method particularly applies to Siberian
XX Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
XX the disease is X-linked. XLPRA1 type is identified in Siberian Huskies,
XX and Samoyeds, while XLPRA2 type is identified in Miniature Schnauzers.
XX The present cDNA sequence is the normal open reading frame (ORF) 15
XX encoding Canine retinitis pigmentosa GTPase regulator (RPGR) protein
XX found in normal dogs.
XX
XX Sequence 2805 BP; 1029 A; 287 C; 940 G; 549 T; 0 other:
SQ
Query Match 30.6%; Score 68; DB 22; Length 2805;
Best Local Similarity 62.2%; Pred. No. 5.9e-07;
Matches 107; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
QY 48 agcagatcaggtcaggtctctgaggaagaaggaaggaaggaaggaagga 107
DB 1071 agggagaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 1130
QY 108 agggagaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag 167
DB 1131 agggagaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 1190
QY 168 caactcgaagcagcagcagcagcagcagcagcagcagcagcagcagcag 219
DB 1191 ggaagggagcaaggaaggaaggaaggaaggaaggaaggaaggaagga 1242
```

Query Match	39.2%;	Score 87;	DB 21;	Length 87;
Best Local Similarity	100.0%;	Pred. No. 1.7e-11;		
Matches	87;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

Qy	133	gaggaagaggaagaggaaggaag	159
Db	61	gaggaagaggaaggaaggaaggaag	87

RESULT	6
AAD07131	
ID	AAD07131 standard; cDNA; 2800 BP.

AC	AAD07131;
XX	
DT	06-AUG-2001 (first entry)

Canine retinitis pigmentosa GTPase regulator (RPGR) mutant #1 cDNA.

KW Dog; X-linked progressive retinal atrophy 1; XLPR1; genetic marker;
 KW retinitis pigmentosa GTPase regulator; RPRG; Siberian Husky; Samoyed
 KW Miniature Schnauzer; mutant; mteuin; ss.

OS Canis familiaris.
OS Synthetic.

AA		Location/Qualifiers
FH	Key	
FT	CDS	1..882

```

FT      /product= "Canine retinitis pigmentosa GTPase
FT      regulator mutant"
FT      /note= "CDS does not include start codon"

```

```

FT      /partial
FT      replace (877..878, AGAGAA)
FT      mutation
FT      /tag= b
FT      /note= "This deletion replaces Arg with Ile and results
FT      in a premature stop codon"

```

PN WO200138578-A1.

PD 31-MAY-2001.

PF 21-NOV-2000; 2000WO-US31940.

PR 24-NOV-1999; 99US-0167365.

...

PA (CORR) CORNELL RES FOUND INC.
XX
PI Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ
XX
WPI: 2001-367707/38.
DR P-PSDB; AAE02398.

PT Identifying dogs with or carrying X-linked progressive retinal atrophy
PT by detecting retinitis pigmentosa GTPase regulator gene mutation,
XX useful when breeding Husky, Samoyed and Miniature Schnauzer -
CS Claim 73: Page 32-33; 88pp; English.

Claim 73; Page 32-33; 88pp; English.

CC The invention relates to a method for identifying dogs which are
CC genetically normal, are carriers of, or are affected with X-linked
CC progressive retinal atrophy (XLRPA), by testing a biological sample with
CC genetic markers that co-segregate with a XLRPA gene locus. The invention
CC also relates to canine retinitis pigmentosa GTPase regulator (RPGR),
CC mutants and their corresponding nucleic acid molecules. The mutated RPGR
CC genes are responsible for the XLRPA in dogs. Methods are used to select
CC dogs for breeding so that dogs carrying the mutated locus are eliminated
CC from the breeding stock. The method particularly applies to Siberian
CC Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
CC the disease is X-linked. XLRPA type is identified in Siberian Huskies,
CC and Samoyeds, while XLRPA2 type is identified in Miniature Schnauzers.
CC The present cDNA sequence is the mutant open reading frame (ORF) 15
CC encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant
CC found in XLRPA1-affected dogs. This mutant is obtained by deleting
CC 'AGAG' nucleotide bases from position 878 to 882 of the wild-type
CC canine RPGR cDNA.

Sequence 2800 BP; 1026 A; 287 C; 938 G; 549 T; 0 other,

Query Match	30.6%	Score 68;	DB 22;	Length 2800;
Best Local Similarity	62.2%;	Pred. No. 5.9e-07;		
Matches 107; Conservative	0;	Mismatches 65;	Indels 0;	Gaps 0

DY 48 agcagaatcacgctcaggttcttgaaagaabaggaaggagaaggaaaggaa 107
 ||| |
Db 1065 aacggagaaaagaaagaaacatttttgcataaacctttaaagtcggcggaaga 1120
 ||| |

Dy 108 agggagcaccagtgaagaatcagagygaggaaaggaagagagagagacgcgacag 167
||||| |||||
Dd 1126 agcgacaqaacccgaagcttagaacqaaaagcaagcaatcraaqagaaacaaacaaacaaacaa 1185

DY 168 caactctgagggcatcagagcagttctgccgaagaagtaagtgcyyaagaa 219
| | | | | | | | | | | | | | |
Dh 1186 ggaatataaaccaatataacataaaaacgaataaaaaaacccatataacaa 172

7 RESULTS

AAD07132
ID AAD07132 standard; cDNA; 2803 BP.
XY

AC	AAD0/132;
XX	
DT	06-AUG-2001 (first entry)
XX	

Canine retinitis pigmentosa GTPase regulator (RPGR) mutant #2 cDNA.

KW Dog; X-linked progressive retinal atrophy 2; XLPR2; genetic marker;
 KW reinitia pigmentosa (GRPase regulator); RPRG; Siberian Husky; Samoyed;
 Miniture Schnauzer. mutant. mutant. ss

XX
XX
Canis familiaris

05 Synthetic.

EH	Key	Location/Qualifiers
ET	CDS	1-1149

```

/* tag = 3
/ product =
# Canine related references CMBac

```

Key

xx	PA	(VLA-)	VLAAMS	INTERUNIVERSITAIR	INST	BIOTECHNOG.	
xx	PI	Cornelis S,	Beyaert R;				
xx	DR	WP1; 2000-499331/44.					
xx	PT	Nucleic acids encoding internal ribosome entry sequences useful for					
xx	PT	directing protein expression in gene therapy procedures -					
PS		Disclosure: Page 32-33; 57pp; English.					
xx	CC	The PITSLRE protein kinase gene can be translated to give two isoforms,					
xx	CC	p110 and p58. Transcription of p58 mRNA was found to start at an					
xx	CC	internal ribosome entry site (IRES). The IRES element was found to be					
xx	CC	cell cycle regulated, p58 is produced during the G2/M stage of the cell					
xx	CC	cycle. The present sequence is the human PITSLRE protein kinase isoform					
xx	CC	alpha2-2. This sequence contains the IRES. The IRES sequence and					
xx	CC	fragments of it may be used to induce cell cycle dependent initiation					
xx	CC	of translation in eukaryotic cells. Vectors containing the IRES may be					
xx	CC	used for the preparation of compositions for the treatment of and/or					
xx	CC	prevention of a disease by gene therapy. Such diseases may be cancer					
xx	CC	and restenosis.					
SQ		Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other:					
	Query Match	100.0%; Score 222; DB 21; Length 2471;					
	Best Local Similarity	100.0%; Pred. No. 5.2e-42;					
	Matches 222; Conservative	0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	gacatcagcgacagcgagaggaagaccagctcgccgagctctctgctcagcaatcagc 60					
DB	907	gacatcagcgacagcgagaggaagaccagctcgccgagctctctcagcaatcagc 966					
QY	61	tcaaggtctcgagaaagagagagagagagagagagagagagagagagagagccacagt 120					
DB	967	tcaaggtctcgagaaagcaccagt 1026					
QY	121	gaagaatcag 180					
DB	1027	gaagaatcag 1086					
QY	181	gcatcagagcagctctcccgaaagaagtaagtgaaggagaagatg 222					
DB	1087	gcatcagagcagctctcccgaaagaagtaagtgaaggagaagatg 1128					
RESULT	5						
AAA73715							
ID	AAA73715	standard; DNA; 87 BP.					
AC	AAA73715;						
DT	07-DEC-2000	(first entry)					
XX							
DE		Functional fragment of human PITSLRE protein kinase gene IRES.					
XX							
KW		IREs; internal ribosome entry site; PITSLRE; human; gene therapy;					
KW		cancer; restenosis; p58; p110; protein kinase; ds.					
XX							
OS		Homo sapiens.					
XX							
PN		WO200044896-A1.					
PD		03-AUG-2000.					
XX							
PF		26-JAN-2000; 2000MO-EP00643.					
XX							
PR		26-JAN-1999; 99EP-0200216.					
XX							
PA		(VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.					
XX							

XX Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
XX Claim 4; Page 32; 57pp; English.
XX
CC The PITSURE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the PITSURE protein kinase IRES. The
CC IRES sequence and fragments of it may be used to induce cell cycle
CC dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.
XX
XX
SQ Sequence 222 BP; 77 A; 36 C; 89 G; 20 T; 0 other;

Query Match 100.0%; Score 222; DB 21; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.7e-42;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacgaggaagagaccagctcgccgagtcctctgacgacgaatcagc 60
|||||
Db 1 gacatcagcgacgaggaagagaccagctcgccgagtcctctgacgacgaatcagc 60
QY 61 tcaggtcttgaggaagagagagaggaagaggaagaggaagaggaagaggaagc 120
|||||
Db 61 tcaggtcttgaggaagagagagaggaagaggaagaggaagaggaagaggaagc 120
QY 121 gaagaaatcagagaggaagagagagagagagagagagagagagagagagagag 180
|||||
Db 121 gaagaaatcagagaggaagagagagagagagagagagagagagagagagagag 180
QY 181 gcacacagagcagctcgccgaagaagtaagtgaagaagaatg 222
|||||
Db 181 gcacacagagcagctcgccgaagaagtaagtgaagaagaatg 222

RESULT 2
AA73710
ID AA73710 standard; RNA; 222 BP.

XX AC AA73710;

XX DT 07-DEC-2000 (first entry)

DE RNA of human PITSURE protein kinase gene internal ribosome entry site.

KW IRES; internal ribosome entry site; PITSURE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT misc_signal 1..222
FT /*tag= a
FT /label= IRES
FT /note= "Internal ribosome entry site"

XX PN WO200044896-A1.

XX PD 03-AUG-2000.

XX PF 26-JAN-2000; 2000WO-EP00643.

XX PR 26-JAN-1999; 99EP-0200216.

XX PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX PI Cornelis S, Beyaert R;

XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
XX
XX Disclosure; Page 32; 57pp; English.
XX
CC The PITSURE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the corresponding RNA of the PITSURE
CC IRES described in AA73710. The IRES sequence and fragments of it may be
CC used to induce cell cycle dependent initiation of translation in
CC eukaryotic cells. Vectors containing the IRES may be used for the
CC preparation of compositions for the treatment of and/or prevention of
CC a disease by gene therapy. Such diseases may be cancer and restenosis.
XX
XX
SQ Sequence 222 BP; 77 A; 36 C; 89 G; 20 U; 0 other;

Query Match 100.0%; Score 222; DB 21; Length 222;
Best Local Similarity 91.0%; Pred. No. 3.7e-42;
Matches 202; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacgaggaagagaccagctcgccgagtcctctgtaagaaatcagc 60
|||||
Db 1 gacatcagcgacgaggaagagaccagctcgccgagtcctctgtaagaaatcagc 60
QY 61 tcaggtcttgaggaagagagagaggaagaggaagaggaagaggaagaggaagc 120
|||||
Db 61 ucagguucugaggaagagagagagaggaagaggaagaggaagaggaagaggaagc 120
QY 121 gaagaaatcagagaggaagagagagagagagagagagagagagagagagagag 180
|||||
Db 121 gaagaaatcagagaggaagagagagagagagagagagagagagagagagagag 180
QY 181 gcacacagcagctcgccgaagaagtaagtgaagaagaatg 222
|||||
Db 181 gcaucaagagcagucuccgcaagaagaaugagagaagaatg 222

RESULT 3

AA73714
ID AA73714 standard; DNA; 660 BP.

XX AC AA73714;

XX DT 07-DEC-2000 (first entry)

DE Human PITSURE protein kinase gene IRES and upstream sequence.

KW IRES; internal ribosome entry site; PITSURE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT misc_signal 439..660
FT /*tag= a
FT /label= IRES
FT /note= "Internal ribosome entry site"

XX PN WO200044896-A1.

XX PD 03-AUG-2000.

XX PF 26-JAN-2000; 2000WO-EP00643.

XX PR 26-JAN-1999; 99EP-0200216.

XX PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 17:07:19 ; Search time 355.05 Seconds
(without alignments)
536.055 Million cell updates/sec

Title: US-09-915-060-1

Perfect score: 222

Sequence: 1 gacatcgcgcagcagcagag.....aagtaagtgcgaagaatg 222

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101: *
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	222	100.0	222	21	AAA73709 Human PITSURE prot
2	222	100.0	222	21	AAA73710 RNA of human PITSURE
3	222	100.0	660	21	AAA73714 Human PITSURE prot
4	222	100.0	2471	21	AAA73711 Human PITSURE prot
5	87	39.2	87	21	AAA73715 Functional fragmen
6	68	30.6	2800	22	AD07131 Canine retinitis p
7	68	30.6	2803	22	AD07132 Canine retinitis p
8	68	30.6	2805	22	AD07130 Canine retinitis p
9	66.8	30.1	49999	20	AA23881 Murine LOBO genom
10	66.8	30.1	49999	20	AA23896 Murine LOBO genom
11	65.4	29.5	49999	20	AA23901 Human LOBO homolo

12	65	29.3	15672	12	AA01613
13	64.6	29.1	2643	14	AA039212
14	64.6	29.1	3717	21	AAA64660
15	64.2	28.9	470	22	AA115222
16	64.2	28.9	470	22	AA133875
17	64.2	28.9	470	22	AA102430
18	62.6	28.2	575	22	AA123315
19	62.6	28.2	575	22	AA148636
20	62.6	28.2	575	22	AA108956
21	62.6	28.2	1969	22	AA114105
22	62.6	28.2	1969	22	AA135486
23	62.6	28.2	1969	22	AA103958
24	62.6	28.2	51259	18	AA083007
25	62	27.9	349	22	AA116341
26	62	27.9	349	22	AA136341
27	62	27.9	349	22	AA104761
28	61.6	27.7	3399	17	AA05868
29	60.6	27.3	3211	18	AA089346
30	60.6	27.3	3901	18	AA089345
31	60.4	27.2	209	22	AA124110
32	60.4	27.2	209	22	AA149402
33	60.4	27.2	209	22	AA109683
34	59.8	26.9	299	22	AA124183
35	59.8	26.9	299	22	AA149470
36	59.8	26.9	299	22	AA109746
37	58.4	26.3	438	22	AA136482
38	58.4	26.3	438	22	AA104897
39	58.4	26.3	117213	19	AA062176
40	58.2	26.2	276	22	AA123164
41	58.2	26.2	2784	18	AA065102
42	58	26.1	475	22	AA11457
43	58	26.1	475	22	AA132728
44	58	26.1	475	22	AA101373
45	58	26.1	512	22	AA106370

ALIGNMENTS

RESULT 1	
AAA73709	standard; DNA; 222 BP.
ID	AAA73709
XX	AAA73709;
AC	07-DEC-2000 (first entry)
DT	
XX	
DE	Human PITSURE protein kinase gene internal ribosome entry site.
XX	
KW	IRRS; internal ribosome entry site; PITSURE; human; gene therapy;
KW	cancer; retinosis; p58; p110; protein kinase; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	misc_signal
FT	1..222
FT	/*tag= a
FT	/label= IRES
FT	/note= "internal ribosome entry site"
XX	
PN	WO200044896-A1.
XX	
PD	03-AUG-2000.
XX	
PF	26-JAN-2000; 2000WO-EP00643.
XX	
PR	26-JAN-1999; 99EP-0200216.
XX	
PA	(VIAA-) VIAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX	
PI	Cornelis S, Beyaert R;
XX	
DR	WPI; 2000-499331/44.

Rianodin receptor
CENP-B CDNA. Homo
DNA encoding cent
Probe #2455 for ge
Probe #2521 used t
Probe #4241 used t
Probe #13248 for g
Probe #17322 used
Probe #8947 used t
Probe #4038 for ge
Probe #4172 used t
Probe #3949 used t
Partial mouse WRN
Probe #4928 for ge
Probe #5027 used t
Probe #4752 used t
Chicken leucocytos
Human p160 CDNA 16
Human p160 CDNA 16
Probe #14043 for g
Probe #18088 used
Probe #674 used t
Probe #14116 for g
Probe #18156 used
Probe #9737 used t
Probe #5168 used t
Probe #4888 used t
HSV-2 str19 SB5 C
Probe #13097 for g
Luciferase gene un
Probe #1390 for ge
Probe #1414 used t
Probe #1364 used t
Probe #6361 used t

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 66.7%; Score 20; DB 21; Length 831;
Best Local Similarity 82.1%; Pred. No. 66;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 2 acatcagcgacagcgagaggaagaccag 29
||||| | ||||| ||| |||||
Db 53 acatcacacacagcaagaacagaccag 80

Search completed: December 4, 2001, 19:03:36
Job time: 6977 sec


```

XX Erythropoietin; Diamond Blackfan anaemia; polycythemia vera; ss.
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX sig_peptide 28..99
XX mat_peptide /tag= a
XX polyA_site 100..1548
XX /tag= b
XX /tag= c
XX 1712..1740
XX
XX WO9008822-A.
XX
XX 09-AUG-1990.
XX
XX 01-FEB-1990; 90WO-US00635.
XX
XX 03-FEB-1989; 89US-0306503.
XX
XX (GENE-) GENETICS INST INC.
XX (WHIT-) WHITEHEAD INST.
XX
XX D'andrea A, Wong G;
XX
XX WPI; 1990-260931/34.
XX P-PSDB; AAR06511.
XX
XX Erythropoietin receptor and gene - used for developing reagents
XX and systems to control and study erythropoiesis.
XX
XX Claim 7; Fig 1; 53pp; English.
XX
XX The sequence was obtd. from a clone isolated from a cDNA library
XX prepd. from uninduced murine erythroleukemia (MEL) cells. The
XX sequence encodes a type I transmembrane protein with binding
XX affinity for EPO. The gene and recombinant EPO receptor produced
XX on expression of the DNA are used to develop reagents and systems
XX to control and study erythropoiesis. It is believed that the EPO
XX receptor is dysfunctional in individuals with Diamond Blackfan
XX anaemia, and may be hyperactive in polycythemia vera.
XX See also AA05748 (human EPO receptor clone).
XX
XX Sequence 1740 BP; 352 A; 527 C; 489 G; 372 T; 0 other;
SQ

```

Query Match 68.7%; Score 20.6; DB 11; Length 1740;
 Best Local Similarity 85.2%; Pred. No. 42;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 catcagcgacagcgaggaagaccag 29
 1 11111 111111111 1111111

DB 825 CGTCAGCAACAGCGAGTGTGAGACCCAG 799

RESULT 12
 AA053994/C
 ID AA053994 standard; CDNA; 1741 BP.
 XX
 XX AA053994;
 XX
 XX 24-JUN-1994 (first entry)
 XX
 XX MEL EPO receptor cDNA.
 XX
 XX Erythropoietin receptor; recombinant; murine; anaemia; ss.
 XX
 XX Mus musculus.
 XX
 XX Key Location/Qualifiers
 XX CDS 28..1551
 XX /tag= a

```

FT sig_peptide 28..100
FT /tag= b
XX
XX US5278065-A.
XX
XX 11-JAN-1994.
XX
XX 03-FEB-1989; 89US-0306503.
XX
XX 03-FEB-1989; 89US-0306503.
XX
XX 25-MAR-1991; 91US-0678877.
XX
XX (CHIT-) CHILDRENS MEDICAL CENT.
XX (GEMV ) GENETICS INST INC.
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
XX D'andrea A, Jones SS, Wong GG;
XX
XX WPI; 1994-025409/03.
XX P-PSDB; AAR47517.
XX
XX Recombinant DNA encoding erythropoietin receptor - used to
XX develop prods. for study, treatment or diagnosis of disorders in
XX which receptor is dysfunctional
XX
XX Claim 1; Fig 2; 24pp; English.
XX
XX Mouse erythroleukemia (MEL) cells were used to construct a cDNA
XX library. The cDNA was used to transfect COS-1 cells and these were
XX screened for radiolodinated erythropoietin (EPO) binding to isolate
XX cDNA encoding the EPO receptor. The cDNA may be used to isolate the
XX EPO receptor from other sources and to study, treat or diagnose
XX disorders in which the EPO receptor is dysfunctional.
XX See also AA053995.
XX
XX Sequence 1741 BP; 353 A; 528 C; 488 G; 372 T; 0 other;
SQ

```

Query Match 68.7%; Score 20.6; DB 15; Length 1741;
 Best Local Similarity 85.2%; Pred. No. 42;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 catcagcgacagcgaggaagaccag 29
 1 11111 111111111 1111111

DB 825 CGTCAGCAACAGCGAGTGTGAGACCCAG 799

RESULT 13
 AA044854/C
 ID AA044854 standard; CDNA; 1741 BP.
 XX
 XX AA044854;
 XX
 XX 19-OCT-1994 (first entry)
 XX
 XX Mouse soluble EPO receptor protein gene.
 XX
 XX Murine; soluble; erythropoietin; EPO; receptor protein; sEPO-R; drug;
 XX antigen; diagnostic agent; biochemical reagent; ss.
 XX
 XX Mus musculus.
 XX
 XX Key Location/Qualifiers
 XX CDS 28..1548
 XX /tag= a
 XX /product= EPO-R
 XX
 XX JP06038787-A.
 XX
 XX 15-FEB-1994.
 XX
 XX 04-MAR-1992; 92JP-0082865.

CC diseases may be cancer and restenosis. The present sequence is the
CC antisense primer di-4 used in creation of a PITSURE distronic vector.
CC This was used in the process of identifying the IRRES.

XX Sequence 35 BP; 3 A; 12 C; 8 G; 12 T; 0 other;

Query Match 93.3%; Score 28; DB 21; Length 35;

Best Local Similarity 100.0%; Pred. No. 0.041;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagacgagagaagacca 28

Db 35 GACATCAGCGACGACGAGAGAACCA 8

RESULT 9

AA118242

ID AA118242 standard; DNA; 522 BP.

AC AA118242;

DT 12-OCT-2001 (first entry)

DE Probe #8175 for gene expression analysis in human cervical cell sample.

KW Probe: human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer; ss.

OS Homo sapiens.

PN W0200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -

PS Claim 25; SEQ ID NO 8175; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human Hela cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX Sequence 522 BP; 107 A; 145 C; 161 G; 109 T; 0 other;

Query Match 68.7%; Score 20.6; DB 22; Length 522;
Best Local Similarity 85.2%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gacatcagcagacgagagaagacc 27

Db 312 gacatcagacagcgtgtgaggaagccc 338

RESULT 10

AA143284

ID AA143284 standard; DNA; 522 BP.

AC AA143284;

DT 17-OCT-2001 (first entry)

DE Probe #11970 used to measure gene expression in human placenta sample.

KW Probe; microarray; human; placenta; antenatal diagnosis;

KW genetic disorder; ss.

OS Homo sapiens.

PN W0200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -

PS Claim 25; SEQ ID NO 11970; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.

XX Sequence 522 BP; 107 A; 145 C; 161 G; 109 T; 0 other;

Query Match 68.7%; Score 20.6; DB 22; Length 522;

Best Local Similarity 85.2%; Pred. No. 37;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gacatcagcagacgagagaagacc 27

Db 312 gacatcagacagcgtgtgaggaagccc 338

RESULT 11

AA005747/C

ID AA005747 standard; cDNA; 1740 BP.

AC AA005747;

DT 04-JAN-1991 (first entry)

DE EPO receptor clone 190.

```

PF 26-JAN-2000; 2000WO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
XX Claim 10; Page 34; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence contains the PITSLRE protein kinase
XX IRES and upstream nucleotides. The IRES sequence and fragments of it
XX may be used to induce cell cycle dependent initiation of translation in
XX eukaryotic cells. Vectors containing the IRES may be used for the
XX preparation of compositions for the treatment of and/or prevention of
XX a disease by gene therapy. Such diseases may be cancer and restenosis.
XX
XX Sequence 660 BP; 210 A; 128 C; 261 G; 61 T; 0 other;

Query Match          100.0%; Score 30; DB 21;
Best Local Similarity 100.0%; Pred. No. 0.0095;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacagcgaggaagaccagc 30
   |||||
DB 439 gacatcagcgacagcgaggaagaccagc 468

RESULT 7
AAA73711
ID AAA73711 standard; DNA; 2471 BP.
XX
XX AAA73711;
XX
XX 07-DEC-2000 (first entry)
XX
XX Human PITSLRE protein kinase isoform alpha2-2.
XX
XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.
XX
XX Homo sapiens.
XX
XX Key          Location/Qualifiers
XX misc_signal 907..1128
XX              /tag= a
XX              /label= IRES
XX              /note= "internal ribosome entry site"
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000WO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX

```

```

PT Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
XX Disclosure; Page 32-33; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence is the human PITSLRE protein kinase isoform
XX alpha2-2. This sequence contains the IRES. The IRES sequence and
XX fragments of it may be used to induce cell cycle dependent initiation
XX of translation in eukaryotic cells. Vectors containing the IRES may be
XX used for the preparation of compositions for the treatment of and/or
XX prevention of a disease by gene therapy. Such diseases may be cancer
XX and restenosis.
XX
XX Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other;

Query Match          100.0%; Score 30; DB 21;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacagcgaggaagaccagc 30
   |||||
DB 907 gacatcagcgacagcgaggaagaccagc 936

RESULT 8
AAA73726/C
ID AAA73726 standard; DNA; 35 BP.
XX
XX AAA73726;
XX
XX 07-DEC-2000 (first entry)
XX
XX Antisense primer di-4 used in creation of PITSLRE dicistronic vector.
XX
XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; antisense; PCR primer;
XX ss.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000WO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
XX Examples; Page 25; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The IRES sequence and fragments of it may be used to induce cell
XX cycle dependent initiation of translation in eukaryotic cells. Vectors
XX containing the IRES may be used for the preparation of compositions for
XX the treatment of and/or prevention of a disease by gene therapy. Such
XX

```

```
RESULT 4
AAA73710
ID AAA73710 standard; RNA; 222 BP.
XX
XX AAA73710;
XX
XX 07-DEC-2000 (first entry)
XX
XX RNA of human PITSLRE protein kinase gene internal ribosome entry site.
XX
XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_signal 1..222
XX /*tag= a
XX /label= IRES
XX /note= "Internal ribosome entry site"
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000MO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VIAAMS INTERUNIVERSITAIR INST BIOTECHNOC.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
XX Disclosure; Page 32; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence is the corresponding RNA of the PITSLRE
XX IRES described in AAA73710. The IRES sequence and fragments of it may be
XX used to induce cell cycle dependent initiation of translation in
XX eukaryotic cells. Vectors containing the IRES may be used for the
XX preparation of compositions for the treatment of and/or prevention of
XX a disease by gene therapy. Such diseases may be cancer and restenosis.
XX
XX Sequence 222 BP; 77 A; 36 C; 89 G; 20 U; 0 other:
XX
XX Query Match 100.0%; Score 30; DB 21; Length 222;
XX Best Local Similarity 96.7%; Pred. No. 0.0085;
XX Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 gacatcagcagcagcgaggaagaccagc 30
XX ||||:|||||||||||||||||||||||
XX Db 1 gacaucagcagcagcgaggaagaccagc 30
XX
XX RESULT 5
XX AAA73713
XX ID AAA73713 standard; DNA; 468 BP.
XX
XX AAA73713;
XX
XX 07-DEC-2000 (first entry)
XX
XX 5' extension of PITSLRE protein kinase internal ribosome entry site.
XX
```

```
KW IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.
XX
XX Homo sapiens.
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000MO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VIAAMS INTERUNIVERSITAIR INST BIOTECHNOC.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
XX Claim 9; Page 33; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence contains a 5' fragment of the PITSLRE
XX protein kinase IRES and nucleotides upstream of the IRES. The IRES
XX sequence and fragments of it may be used to induce cell cycle
XX dependent initiation of translation in eukaryotic cells. Vectors
XX containing the IRES may be used for the preparation of compositions for
XX the treatment of and/or prevention of a disease by gene therapy. Such
XX diseases may be cancer and restenosis.
XX
XX Sequence 468 BP; 144 A; 100 C; 182 G; 42 T; 0 other:
XX
XX Query Match 100.0%; Score 30; DB 21; Length 468;
XX Best Local Similarity 100.0%; Pred. No. 0.0091;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 gacatcagcagcagcgaggaagaccagc 30
XX |||||:|||||||||||||||||||||
XX Db 439 gacatcagcagcagcgaggaagaccagc 468
XX
XX RESULT 6
XX AAA73714
XX ID AAA73714 standard; DNA; 660 BP.
XX
XX AAA73714;
XX
XX 07-DEC-2000 (first entry)
XX
XX Human PITSLRE protein kinase gene IRES and upstream sequence.
XX
XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_signal 439..660
XX /*tag= a
XX /label= IRES
XX /note= "Internal ribosome entry site"
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
```

CC The PITSLRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is a 5' fragment of PITSLRE protein kinase
CC IRES. The IRES sequence and fragments of it may be used to induce cell
CC cycle dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.

XX
SQ Sequence 30 BP; 11 A; 8 C; 10 G; 1 T; 0 other;

Query Match 100.0%; Score 30; DB 21; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.0069;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcagcagcagcagcagcagcagcagc 30
|||||
DB 1 gacatcagcagcagcagcagcagcagcagc 30

RESULT 2

AA73728
ID AAA73728 standard; DNA; 40 BP.

XX
AC AAA73728;

XX
DT 07-DEC-2000 (first entry)

DE Sense primer di-6 used in creation of PITSLRE dicistronic vector.

XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;

KM cancer; restenosis; p58; p110; protein kinase; PCR primer;

XX

OS Homo sapiens.

OS Synthetic.

PN WO200044896-A1.

XX
PD 03-AUG-2000.

XX
PF 26-JAN-2000; 2000WO-EP00643.

XX
PR 26-JAN-1999; 99EP-0200216.

XX
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX
PI Cornelis S, Beyaert R;

XX
DR WPI: 2000-499331/44.

XX
PT Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -

XX
PS Examples: Page 25; 57pp; English.

XX
CC The PITSLRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The IRES sequence and fragments of it may be used to induce cell
CC cycle dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis. The present sequence is the
CC sense primer di-6 used in creation of a PITSLRE dicistronic vector.
CC This was used in the process of identifying the IRES.

XX
SQ Sequence 40 BP; 14 A; 10 C; 12 G; 4 T; 0 other;

Query Match 100.0%; Score 30; DB 21; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcagcagcagcagcagcagcagcagc 30
|||||
DB 11 gacatcagcagcagcagcagcagcagcagc 40

RESULT 3

AA73709
ID AAA73709 standard; DNA; 222 BP.

XX
AC AAA73709;

XX
DT 07-DEC-2000 (first entry)

DE Human PITSLRE protein kinase gene internal ribosome entry site.

XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;

KM cancer; restenosis; p58; p110; protein kinase; ds.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT misc_signal 1..222

FT /*tag= a

FT /label= IRES

FT /note= "internal ribosome entry site"

XX
PN WO200044896-A1.

XX
PD 03-AUG-2000.

XX
PF 26-JAN-2000; 2000WO-EP00643.

XX
PR 26-JAN-1999; 99EP-0200216.

XX
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX
PI Cornelis S, Beyaert R;

XX
DR WPI: 2000-499331/44.

XX
PT Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -

XX
PS Claim 4; Page 32; 57pp; English.

XX
CC The PITSLRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the PITSLRE protein kinase IRES. The
CC IRES sequence and fragments of it may be used to induce cell cycle
CC dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.

XX
SQ Sequence 222 BP; 77 A; 36 C; 89 G; 20 T; 0 other;

Query Match 100.0%; Score 30; DB 21; Length 222;

Best Local Similarity 100.0%; Pred. No. 0.0085;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcagcagcagcagcagcagcagcagc 30
|||||
DB 1 gacatcagcagcagcagcagcagcagcagc 30

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: December 4, 2001, 19:03:35 ; Search time 355.05 Seconds
(without alignments)
72.440 Million cell updates/sec

Title: US-09-915-060-4
Perfect score: 30
Sequence: 1 gacatcagcgacagcgagagagacagc 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_1101: *
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	21	AAAT3712
2	30	100.0	40	21	AAAT3728
3	30	100.0	222	21	AAAT3709
4	30	100.0	222	21	AAAT3710
5	30	100.0	468	21	AAAT3713
6	30	100.0	660	21	AAAT3714
7	30	100.0	2471	21	AAAT3711
8	28	93.3	35	21	AAAT3726
9	20.6	68.7	522	22	AA118242
10	20.6	68.7	522	22	AA143284
11	20.6	68.7	1740	11	AA05747

C 12	20.6	68.7	1741	15	AA053994
C 13	20.6	68.7	1741	15	AA044854
C 14	20.6	68.7	1741	15	AA081891
C 15	20	66.7	831	21	AA049954
C 16	20	66.7	833	21	AA034562
C 17	20	66.7	1722	19	AAV20740
C 18	20	66.7	2443	19	AAV36468
C 19	20	66.7	2450	19	AAV20757
C 20	20	66.7	2918	21	AAZ43922
C 21	20	66.7	2919	21	AAZ24486
C 22	19.6	65.3	1261	21	AA067837
C 23	19.4	64.7	292	19	AAV62323
C 24	19.4	64.7	601	19	AAV40880
C 25	19.4	64.7	601	21	AAAG6448
C 26	19.4	64.7	601	22	AAV98445
C 27	19.4	64.7	1128	21	AAAT30619
C 28	19.4	64.7	1128	21	AAAT30629
C 29	19.4	64.7	1128	21	AAAT30724
C 30	19.4	64.7	1128	21	AAAT30728
C 31	19.4	64.7	1648	21	AAAT11700
C 32	19.4	64.7	1841	19	AAV62310
C 33	19.4	64.7	2582	17	AAAT36034
C 34	19.4	64.7	2582	20	AAAT26245
C 35	19.4	64.7	2582	20	AAV81862
C 36	19.4	64.7	2582	21	AAAT88577
C 37	19.4	64.7	2582	21	AAZ89790
C 38	19.4	64.7	2582	21	AAZ88006
C 39	19.4	64.7	2776	18	AAAT73366
C 40	19.4	64.7	2894	19	AAV62334
C 41	19	63.3	182	14	AAQ42773
C 42	19	63.3	1077	21	AAAT21043
C 43	19	63.3	1077	21	AAAT34921
C 44	19	63.3	2296	19	AAV12457
C 45	19	63.3	2372	21	AAAT21041

ALIGNMENTS

RESULT 1	
ID AAA73712 standard; DNA; 30 BP.	
XX	
AC AAA73712;	
XX	
DT 07-DEC-2000 (first entry)	
XX	
DE 5' fragment of PITSURE protein kinase internal ribosome entry site.	
XX	
KM IRES; Internal ribosome entry site; PITSURE; human; gene therapy;	
KW cancer; retinoblastoma; p53; p110; protein kinase; ds.	
XX	
OS Homo sapiens.	
XX	
PN WO200044896-A1.	
XX	
PD 03-AUG-2000.	
XX	
PF 26-JAN-2000; 2000WO-EP00643.	
XX	
PR 26-JAN-1999; 99EP-0200216.	
XX	
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.	
XX	
PI Cornelis S, Beyaert R;	
XX	
DR WPI: 2000-499331/44.	
XX	
PT Nucleic acids encoding internal ribosome entry sequences useful for	
XX	
PS directing protein expression in gene therapy procedures -	
XX	
Claim 8; Page 33; 57pp; English.	

FILE REFERENCE: 30062-20022.00 Polyketide Synthase Enz
CURRENT APPLICATION NUMBER: US/09/154,083
CURRENT FILING DATE: 1998-09-16
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 750
TYPE: DNA
ORGANISM: Streptomyces hygroscopicus
US-09-154-083-11

Query Match 62.7%; Score 18.8; DB 1; Length 750;
Best Local Similarity 76.7%; Pred. No. 50;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gacatcagcgacagcgagaggaagaccagc 30
|||||
DB 491 GACAGCAGCGACGTCGAGCGACGTCGACG 462

RESULT 15
US-08-222-289-1/c
Sequence 1, Application US/08222289
Patent No. 5559010
GENERAL INFORMATION:
APPLICANT: Kilonsky, Daniel J
APPLICANT: Destruelle, Monika
TITLE OF INVENTION: NUTRIENT REGULATED GENE EXPRESSION
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,289
FILING DATE: 04-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59767/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2165 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 370..1434
US-08-222-289-1

Query Match 62.7%; Score 18.8; DB 1; Length 2165;
Best Local Similarity 76.7%; Pred. No. 55;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 gacatcagcgacagcgagaggaagaccagc 30

DB 1002 GTCATCAGCATAGCAACAGGCAACCCAGC 973
|||||

Search completed: December 4, 2001, 18:57:33
Job time: 9199 sec

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,496
FILING DATE: 06-OCT-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,654
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-944-496-6

Query Match 64.7%; Score 19.4; DB 3; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 2 acatcagcagcagcagcagcagcagc 30
Db 809 ACATCAGCTTCCGCGAGAGATGACCATC 837

RESULT 10
US-08-925-767-6
Sequence 6, Application US/08925767
Patent No. 6225084
GENERAL INFORMATION:
APPLICANT: FAIB, DEAN A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/925,767
FILING DATE: 09-SEPT-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844

FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-097
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-925-767-6

Query Match 64.7%; Score 19.4; DB 4; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 2 acatcagcagcagcagcagcagcagc 30
Db 809 ACATCAGCTTCCGCGAGAGATGACCATC 837

RESULT 11
US-08-330-108-5/C
Sequence 5, Application US/08330108
Patent No. 5795752
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A.
TITLE OF INVENTION: IL-2-Stimulated Gene
EXPRESSION
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/330,108
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/104,736
FILING DATE:
APPLICATION NUMBER: US/07/796,066
FILING DATE:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 base pairs
TYPE: nucleic acid
STRANDEDNESS: single stranded
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: T-cell blast
IMMEDIATE SOURCE:
LIBRARY:

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,654
FILING DATE: 09-FEB-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-041
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-08-599-654-6

Query Match 64.7%; Score 19.4; DB 2; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 acatcagcgacgcgaggaagaccagc 30
Db 809 ACATCAGCTCCGCGAGAGATGACCATC 837

RESULT 5
US-08-485-573-6
Sequence 6, Application US/08485573
Patent No. 5968770
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,573
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-032
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-08-485-573-6

Query Match 64.7%; Score 19.4; DB 2; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 acatcagcgacgcgaggaagaccagc 30
Db 809 ACATCAGCTCCGCGAGAGATGACCATC 837

RESULT 6
US-08-944-868A-6
Sequence 6, Application US/08944868A
Patent No. 6018025
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,868A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/599,654
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-041
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-08-944-868A-6

QY 3 catcagcgacagcgaggaagaccagc 30
||| ||||| ||||| |||||
DB 799 CACGAGGCCGACGAGAGAGACACGAGC 772

RESULT 2
US-08-480-994-6
; Sequence 6, Application US/08480994
; Patent No. 5834248
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,994
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-08-480-994-6

Query Match 64.7%; Score 19.4; DB 2; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 acatcagcgacgagcgaggaagaccagc 30
||||| ||||| ||||| |||||
DB 809 ACATCAGCTTCGCGAGAGATGACCATC 837

RESULT 3
US-08-616-844-6
; Sequence 6, Application US/08616844
; Patent No. 5849578
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,844
; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-08-616-844-6

NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,844
FILING DATE: 15-MAR-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,654
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-616-844-6

Query Match 64.7%; Score 19.4; DB 2; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 acatcagcgacgagcgaggaagaccagc 30
||||| ||||| ||||| |||||
DB 809 ACATCAGCTTCGCGAGAGATGACCATC 837

RESULT 4
US-08-599-654-6
; Sequence 6, Application US/08599654
; Patent No. 5882925
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,844
; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-08-616-844-6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:57:28 : Search time 152.61 Seconds
(without alignments)
44.521 Million cell updates/sec

Title: US-09-915-060-4

Perfect score: 30

Sequence: 1 gacatcagcagcagcagcagcagcagcagc 30

Scoring table: IDENTITY_NUC

Searched: Gap0 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	66.7	2443	2	US-08-745-934-2
2	19.4	64.7	2582	2	US-08-480-994-6
3	19.4	64.7	2582	2	US-08-616-844-6
4	19.4	64.7	2582	2	US-08-599-654-6
5	19.4	64.7	2582	2	US-08-485-573-6
6	19.4	64.7	2582	3	US-08-944-868A-6
7	19.4	64.7	2582	3	US-08-944-423A-6
8	19.4	64.7	2582	3	US-08-925-743-6
9	19.4	64.7	2582	3	US-08-944-496-6
10	19.4	64.7	2582	4	US-08-925-767-6
11	19.4	63.3	182	5	US-08-330-108-5
12	19.4	63.3	182	5	US-08-330-108-5
13	19.4	63.3	182	5	US-08-330-108-5
14	19.4	63.3	182	5	US-08-330-108-5
15	19.4	63.3	182	5	US-08-330-108-5
16	19.4	63.3	182	5	US-08-330-108-5
17	19.4	63.3	182	5	US-08-330-108-5
18	19.4	63.3	182	5	US-08-330-108-5
19	19.4	63.3	182	5	US-08-330-108-5
20	19.4	63.3	182	5	US-08-330-108-5
21	19.4	63.3	182	5	US-08-330-108-5
22	19.4	63.3	182	5	US-08-330-108-5
23	19.4	63.3	182	5	US-08-330-108-5
24	19.4	63.3	182	5	US-08-330-108-5
25	19.4	63.3	182	5	US-08-330-108-5
26	19.4	63.3	182	5	US-08-330-108-5
27	19.4	63.3	182	5	US-08-330-108-5

28	17.8	59.3	1549	2	US-08-865-597A-1	Sequence 1, Appl
29	17.8	59.3	2122	4	US-09-029-603-1	Sequence 1, Appl
30	17.8	59.3	7130	4	US-09-036-105-31	Sequence 31, Appl
31	17.8	59.3	9551	1	US-08-056-200-93	Sequence 93, Appl
32	17.8	59.3	9551	1	US-08-056-200-93	Sequence 93, Appl
33	17.8	59.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl
34	17.8	59.3	4411529	4	US-09-103-840A-1	Sequence 1, Appl
35	17.4	58.0	333	2	US-08-634-797-21	Sequence 21, Appl
36	17.4	58.0	1074	2	US-08-463-081B-29	Sequence 29, Appl
37	17.4	58.0	1074	2	US-08-461-379A-29	Sequence 29, Appl
38	17.4	58.0	1074	2	US-08-462-390B-29	Sequence 29, Appl
39	17.4	58.0	1074	3	US-08-463-074B-29	Sequence 29, Appl
40	17.4	58.0	1074	3	US-08-465-585C-29	Sequence 29, Appl
41	17.4	58.0	1074	3	US-08-652-446-29	Sequence 29, Appl
42	17.4	58.0	2450	2	US-08-463-081B-5	Sequence 29, Appl
43	17.4	58.0	2450	2	US-08-461-379A-5	Sequence 5, Appl
44	17.4	58.0	2450	2	US-08-462-390B-5	Sequence 5, Appl
45	17.4	58.0	2450	3	US-08-463-074B-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-745-934-2/c
; Sequence 2, Application US/08745934
; Patent No. 5861496
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: HUMAN SQUALENE EPOXIDASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/745,934
; FILING DATE: Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0151 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO. 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2443 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-08-745-934-2

Query Match 66.7%; Score 20; DB 2; Length 2443;
Best Local Similarity 82.1%; Pred. No. 19;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Query Match 86.0%; Score 25.8; DB 10; Length 216;
Best Local Similarity 93.1%; Pred. No. 14;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 gacatcagcgacagcgagagaaagaccag 29
||||||| |||||||||
Db 86 GACATCAGTCAGCAGCGAGAGAAACCAG 114

Search completed: December 4, 2001, 18:14:19
Job time: 6760 sec

TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome research. 10 (10), 1617-1630 (2000)
MEDLINE	20493374
PUBMED	11042159
AUTHORS	3 (bases 1 to 1000) Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,T., Nishi,K., Kitsuami,T., Tashiro,H., Itoh.M., Sumi.N., Ishii.Y., Nakamura.S., Hazama.M., Nishino.T., Harada.A., Yamamoto.R., Matsunoto.H., Sakaguchi,S., Ikegami,T., Kashida.K., Fujiwaka,E., Inoue,K., Togawa,Y., Izawa.M., Ohtsu.E., Wataniki.M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka.T., Matsuura.S., Kawai,J., Okazaki.Y., Muramatsu.M., Inoue.Y., Kita,A. and Hayashizaki,Y.
TITLE	Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer
JOURNAL	Genome research. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076661
AUTHORS	4 (bases 1 to 1000) The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection <i>Nature</i> 409, 685-690 (2001) 5 (bases 1 to 1000)
TITLE	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono.H., Arai,A., Atakawa,T., Carninci,P., Fukuda,S., Fukushima,Y., Furuno.M., Hanagawa,T., Hara,A., Hayashi.N., Hiramoto,K., Hirota,F., Horl.F., Imotani,K., Ishii,Y., Itoch.M., Izawa.M., Kato.H., Kawai.J., Kotani,Y., Konno,H., Koude.M., Koya.S., Kuhihara,C., Matsuyama,T., Miyazaki,A., Nishi.K., Nomura,K., Numasaki,R., Ohno.M., Okazaki,Y., Okido.T., Owa.C., Salito.H., Salito,R., Sakai.C., Sakai.K., Sano.H., Sasaki,D., Shibata,K., Shidbate,Y., Shinagawa,A., Shiraki,T., Sobabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima.Y., Toya.T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino.M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-jun-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsukuma-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gscc.rken.go.jp], URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216]
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research group in Riken Genomic Sciences Center and Genome science laboratory in RIKEN-Division of Experimental Animal Research at Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAAGAGAACAACATCCAAAGCGCTGTATTTCCTTTTNN 3'], cDNA was prepared by using triehase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA wenthy through one round of normalization to Rot = 5.0 and subtraction to Rot - 20.0. Second strand cDNA was prepared with the primer adptent of sequence[5' GAAGAGAAGACTTGCAAGTAATAAATGAATGCCCCCCCCC 3']. cDNA was cleaved with XhoI and sstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.
FEATURES	Location/Qualifiers 1..1000 /oranism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /db_xref="MGD:MGI:1098660" /db_xref="MGI:1908754" /clone="2010016J05" /sex="male" /tissue_type="small intestine" /clone_id="RIKEN_full-length enriched mouse cDNA library" /dev stage="adult" 64...>1000 /note="putative"
SOURCE	CDS

BASE COUNT	365 a	177 c	334 g	124 t
ORIGIN	/codon_start=1 /protein_id="BAB25574.1" /db_xref="GI:12842371" /translation="MDGEKDSWKYKTIIDELIQEKKRRKKEDEBEKAETIKLNSDDPSDK RDSLEBSELKDHRREITITNSPYRREDSMERGEDDSLIKPEQDSKREKAHNRD EKREKRRRSHSASEGKAHVYDKEREHERRRKRRREQDKARREWRRKREARPH SRRRRLDQERKERERERKLEQOKRQREKEREERAEERKEREERVAHHRTM REESDCKGKWHMSRSLPDRPREFEREEEDNNKPKVEEKVEERLLDLSIDISSEKRT SSAEFSSAGSGSGFEEFEFEFEFEFEFEFEFESESE"			
Query Match	89.3%	Score 26.8;	DB 12;	Length 1000;
Best Local Similarity	93.3%	Pred. No. 7.8;		
Matches	28;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
0y	1	gacatcagcagcagcagcagcagcagcagc 30		
Db	865	GACATCACTGTACAGCGAGCAAAACCAGC 894		
RESULT 15				
AV105652				
LOCUS	AV105652	216 bp	mRNA	EST
DEFINITION	AV105652	Mus musculus	liver C57BL/6J	13-day embryo
ACCESSION	CDNA clone 2510017D12			Mus musculus
VERSION	AV105652			
KEYWORDS	AV105652.1	GI:5253200		
SOURCE	EST.			
ORGANISM	house mouse.			
	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
	1 (bases 1 to 216)			
	Carpinci, P., Shibata, K., Ozawa, Y., Kono, H., Itoh, M., Atzawa, K.,			
	Akehira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara			
	, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Itawa, M., Kawai, J.,			
	Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C.,			
	Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara			
	, Y., Suzuki, H., Suzuki, H., Tateono, M., Tomaru, Y., Tomimaga, N.,			
	Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,			
	Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.			
	RIKEN MOUSE ESTs			
	Unpublished (1999)			
JOURNAL	Genome Science Laboratory			
COMMENT	RIKEN			
	3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan			
	Tel: 81-298-36-9145			
	Fax: 81-298-36-9098			
	Email: genome-res@rtc.riken.go.jp			
	Thermotabilization and thermocactivation of thermostable enzymes by			
	treatose and its application for the synthesis of full length cDNAs			
	(Proc. Natl. Acad. Sci. U.S.A. 95(12):520-524 (1998))			
	Transcriptional sequencing: A method for DNA sequencing using RNA			
	polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))			
	Please visit our web site (http://genome.rtc.riken.go.jp) for			
	further details.			
FEATURES	Location/Qualifiers			
source	1..216			
	/organism="Mus musculus"			
	/strain="C57BL/6J"			
	/db_xref="taxon:10090"			
	/clone="2510017D12"			
	/clone_lib="Mus musculus liver C57BL/6J 13-day embryo"			
	/sex="mixed"			
	/tissue_type="liver"			
	/dev_stage="13-day embryo"			
BASE COUNT	79 a	31 c	81 g	25 t
ORIGIN				

ORIGIN

Query Match 89.3%; Score 26.8; DB 10; Length 649;
Best Local Similarity 93.3%; Pred. No. 7.3;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gacatcagcgacagcgaggaagaccagc 30
||||| ||||| ||||| ||||| |||||
Db 44 GACATCAGTGCAGCAGAGGAGAAACCCAGC 73

RESULT 12

BG298479

LOCUS 900 bp mRNA EST 21-FEB-2001
DEFINITION 602396747P1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4511434 5',
mRNA sequence.

ACCESSION BG298479.1 GI:13063173

VERSION BG298479

KEYWORDS EST.

SOURCE

house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 900)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbsr@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLM10394 row: n column: 11

High quality sequence stop: 676.

Location/Qualifiers

FEATURES

source

1..900

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone_image="4511434"

/clone_lib="NIH-MGC_94"

/tissue_type="retina"

/lab_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."

BASE COUNT 302 a 175 c 300 g 123 t

ORIGIN

Query Match 89.3%; Score 26.8; DB 11; Length 900;
Best Local Similarity 93.3%; Pred. No. 7.7;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gacatcagcgacagcgaggaagaccagc 30
||||| ||||| ||||| ||||| |||||
Db 47 GACATCAGTGCAGCAGAGGAGAAACCCAGC 76

RESULT 13

CNS0431P/C

LOCUS

DEFINITION

CNS0431P 972 bp DNA GSS 18-MAY-2000
Tetradon nigroviridis genome survey sequence PUC-ori end of clone
079D08 of library G from Tetradon nigroviridis, genomic survey
sequence.

ACCESSION AL272842
VERSION AL272842.1 GI:7995084

KEYWORDS

GSS; genome survey sequence.

SOURCE

ORGANISM

Tetradon nigroviridis.

Tetradon nigroviridis.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetradon.

REFERENCE 1 (bases 1 to 972)

AUTHORS

Roest-Collins/H., Jaillon/O., Dasilva/C., Fitzames/C., Fisher/C.,
Bonneau/L., Billault/A., Quetier/F., Saurin/W., Bernot/A. and
Weissenbach/J.

TITLE

Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis

JOURNAL

Unpublished

REFERENCE 2 (bases 1 to 972)

AUTHORS

Roest-Collins/H., Jaillon/O., Dasilva/C., Bonneau/L., Fisher/C.,
Bernot/A., Fitzames/C., Wincker/P., Brothier/P., Quetier/F.,
Saurin/W. and Weissenbach/J.

TITLE

Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence

JOURNAL

Unpublished

REFERENCE 3 (bases 1 to 972)

AUTHORS

Genoscope.

COMMENT

Direct Submission

Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases

This sequence is a single read and was generated as part of a large
scale clone-and-sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetradon>.

FEATURES

source

1..972

/organism="Tetradon nigroviridis"

/db_xref="taxon:99883"

/clone_image="079D08"

/clone_lib="G"

/note="Genoscope sequence ID : C0B6079DB04SP1-end :
PUC-ori"

BASE COUNT 224 a 249 c 220 g 272 t

ORIGIN

7 others

Query Match 89.3%; Score 26.8; DB 13; Length 972;
Best Local Similarity 93.3%; Pred. No. 7.8;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gacatcagcgacagcgaggaagaccagc 30
||||| ||||| ||||| ||||| |||||
Db 696 GACATCAGTGCAGTGCAGAGAGAGCCAGC 667

RESULT 14

AK008283

LOCUS

DEFINITION

AK008283

ACCESSION

AK008283

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1000)

AUTHORS

Carninci/P. and Hayashizaki/Y.

TITLE

High-efficiency full-length cDNA cloning
Methods in enzymology. 303, 19-44 (1999)

JOURNAL

MEDLINE

PUBMED

REFERENCE

2 (bases 1 to 1000)

Carninci/P., Shibata/Y., Hayatsu/N., Sugahara/Y., Shibata/K.,
Itoh/M., Konno/H., Okazaki/Y., Muramatsu/M. and Hayashizaki/Y.

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:490063

Putative full length read
vector to vector length is 183
Seq primer: -28ml3 rev1 ET from Amersham.

FEATURES

source

Location/Qualifiers

1..182

/organism="Mus musculus"

/strain="C57BL6 x DBA"

/db_xref="taxon:10090"

/clone="IMAGE:821783"

/clone_1lb="Beddington mouse embryonic region"

/sex="pooled"

/tissue_type="embryo"

/dev_stage="7.5dpc"

/lab_host="DH12S"

/note="Organ: whole embryo; Vector: PCMV-SPORT; Site_1:
Sali; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Gastrulating embryos were collected at 7.5dpc
from C57BL6 x DBA matings, excluding embryos that had
developed head folds and all extraembryonic tissues.
Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
Referenced in Development 121, 2479-2489 (1995)"

BASE COUNT

66 a 32 c 59 g 25 t

ORIGIN

Query Match

Best Local Similarity 93.3%; Score 26.8; DB 10; Length 182;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gacatcagcgacgagaggaaccagc 30
||||| ||||| ||||| ||||| |||||
Db 88 GACATCAGTCAGCAGCAGGAGAAACACG 117

RESULT 10

AL034747

LOCUS AL034747 372 bp mRNA EST 29-DEC-1999
V6608b53 Beddington mouse dissected endoderm Mus musculus CDNA
clone 528_8K10 5', mRNA sequence.

ACCESSION

AL034747

AL034747.1 GI:6646373

VERSION

EST.

KEYWORDS

house mouse.

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 372)

AUTHORS

Harrison,S.M., Dunwoodie,S.L., Arkell,R.M., Lehrach,H. and
Beddington,R.S.

TITLE

Isolation of novel tissue-specific genes from cDNA libraries
representing the individual tissue constituents of the gastrulating
mouse embryo

JOURNAL

Development 121 (8), 2479-2489 (1995)

MEDLINE

95401865

COMMENT

Contact: Wiles,M., Lehrach,H. and Avner,P.
ERC Mouse Transcript Mapping Consortium
Genoscope - CNS
2, rue Gaston Cremieux, 91000 Evry, France
Email: payner@pasteur.fr

FEATURES

clone available from Ressourcenzentrum, Heubnerweg 6, D-14059
Berlin, Germany. Web site http://www.rzp.de
Seq primer: CCGGTCGGGATTCGCGG;
High quality sequence only submitted.
Vector: pSPORT; site_1: NotI; site_2: Sali;
Cloned unidirectionally.
Dissected endoderm 7.5 days.
Average insert size: 1.2 kb (range: 0.2 - 2.kb).
Location/Qualifiers

BASE COUNT

229 a 111 c 207 g 102 t

ORIGIN

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 649)
Marrin,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marrin M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: primer name ambiguous
High quality sequence stop: 426.
Location/Qualifiers

FEATURES

1..649
Location/Qualifiers
/organism="Mus musculus"
/strain="CD1"
/db_xref="taxon:10090"
/clone="IMAGE:2101724"
/clone_1lb="Rashbass mouse MOV 9 5 optic vesicle"
/tissue_type="optic vesicle and lens placode"
/dev_stage="embryo - 9.5 dpc"
/lab_host="DH10B"
/note="Organ: eye; Vector: pSPORT1; mRNA made from
developing eye tissue. cDNA made by oligo-dT priming with
NotI oligo. Sali adaptor (5'-TCGACCCAGCGTCG-3')
ligated to 5' ends. Size-selected with cDNA size
fractionation resin, average insert size 1.3 kb. Primary
library, non-amplified. Library constructed by Dr. Pen
Rashbass (penr@hu.mrc.ac.uk)."

BASE COUNT

229 a 111 c 207 g 102 t

ORIGIN

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 649)
Marrin,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marrin M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: primer name ambiguous
High quality sequence stop: 426.
Location/Qualifiers

FEATURES

1..649
Location/Qualifiers

BASE COUNT

229 a 111 c 207 g 102 t

ORIGIN

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 649)
Marrin,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marrin M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: primer name ambiguous
High quality sequence stop: 426.
Location/Qualifiers

Query Match 100.0%; Score 30; DB 10; Length 1042;
 Best Local Similarity 100.0%; Pred. No. 0.6;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcagcagcagcagcagcagcagc 30
 |||||||||||||||||||||||||||||||
 Db 413 GACATCAGCGACGACGAGAGACACG 442

RESULT 7
 BE740559 1063 bp mRNA EST 15-SEP-2000
 LOCUS 601595656p1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949693 5',
 DEFINITION mRNA sequence.
 BE740559
 VERSION BE740559.1 GI:10154551
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1063)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
 Plate: LNC814 row: P column: 14
 High quality sequence stop: 741.
 Location/Qualifiers
 1. 1063
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3949693"
 /clone_1lb="NIH_MGC_9"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: ovary; Vector: pOTB7; Site: 1; XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 351 a 203 c 402 g 107 t

ORIGIN

Query Match 100.0%; Score 30; DB 10; Length 1063;
 Best Local Similarity 100.0%; Pred. No. 0.6;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcagcagcagcagcagcagcagc 30
 |||||||||||||||||||||||||||||||
 Db 408 GACATCAGCGACGACGAGAGACACG 437

RESULT 8
 BG107702 1340 bp mRNA EST 30-JAN-2001
 LOCUS 602277878p1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4365474 5',
 DEFINITION mRNA sequence.
 BG107702
 ACCESSION BG107702.1 GI:12601548
 VERSION

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1340)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM10014 row: 1 column: 19
 High quality sequence stop: 667.
 Location/Qualifiers
 1. 1340
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4365474"
 /clone_1lb="NIH_MGC_86"
 /tissue_type="osteosarcoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bone; Vector: pCMV-SPORT6; Site: 1; NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.533 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC library."

BASE COUNT 467 a 272 c 450 g 151 t

ORIGIN

Query Match 100.0%; Score 30; DB 11; Length 1340;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcagcagcagcagcagcagcagc 30
 |||||||||||||||||||||||||||||||
 Db 276 GACATCAGCGACGACGAGAGACACG 305

RESULT 9
 AA413247 182 bp mRNA EST 02-MAY-1997
 LOCUS ve52d12.f1 Beddington mouse embryonic region Mus musculus cDNA
 DEFINITION clone IMAGE:821783 5' similar to gb:M58633 Mouse p58/GTA protein
 kinase mRNA, complete cds (MOUSE);, mRNA sequence.
 AA413247
 ACCESSION AA413247.1 GI:2071810
 VERSION
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 182)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,B.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Thaising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMNI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMNI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

```

ACCESSION      BE742222
VERSION        BE742222.1  GI:10156214
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE         NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL       National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT       Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgabbs-remail.nih.gov
               Tissue Procurement: DCTD/DTP
               CDNA Library Preparation: Ling Hong/Rubin Laboratory
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
               Plate: LLCM520 row: j column: 08
               High quality sequence stop: 702.
               Location/Qualifiers
                 1..723
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone_image="3836647"
                   /clone_lib="NIH-MGC_9"
                   /tissue_type="adenocarcinoma cell line"
                   /lab_host="DH10B (phage-resistant)"
                   /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                   EcoRI; CDNA made by oligo-dT priming. Directionally
                   cloned into EcoRI/XhoI sites using the following 5'
                   adaptor: GGCAAGAG(G). Size-selected >500bp for average
                   insert size 1.8kb. Library constructed by Ling Hong in
                   the laboratory of Gerald M. Rubin (University of
                   California, Berkeley) using ZAP-cDNA synthesis kit
                   (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT      230 a 140 c 284 g 69 t
ORIGIN
```

```

Query Match      100.0%; Score 30; DB 10; Length 723;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gacatcagcagcagcagcagcagcagcagc 30
Db 407 GACATCAGCAGCAGCAGCAGCAGCAGCAGC 436
|||||
```

```

RESULT 5
LOCUS      AL043340      740 bp      mRNA      EST      29-FEB-2000
DEFINITION DKEFp43400923_r1.434 (synonym: htes3) Homo sapiens CDNA clone
ACCESSION  AL043340
VERSION    AL043340.1  GI:5422730
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE       1 (bases 1 to 740)
JOURNAL     Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
COMMENT     EST (Blum, et al.)
             Unpublished (1999)
             Contact: Blum H
             MIPS
             Am Klopferspitz 18a D-82152 Martinsried, Germany
             This is the 5' sequence of the clone insert
             Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
             Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
             sequenced by LMU (Ludwig Maximilians University,
```

Munich/Germany) within the CDNA sequencing consortium of the German Genome Project. No st sequence available. This clone (DKEFp43400923) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

FEATURES
source
1..740
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="DKEFp43400923"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
```

```

BASE COUNT      230 a 155 c 254 g 100 t 1 others
ORIGIN
```

```

Query Match      100.0%; Score 30; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gacatcagcagcagcagcagcagcagcagc 30
Db 44 GACATCAGCAGCAGCAGCAGCAGCAGCAGC 73
|||||
```

```

RESULT 6
LOCUS      BE742943      1042 bp      mRNA      EST      15-SEP-2000
DEFINITION 601574925P1 NIH-MGC_9 Homo sapiens CDNA clone IMAGE:3835686 5',
MRNA sequence.
ACCESSION  BE742943
VERSION    BE742943.1  GI:10157033
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE       1 (bases 1 to 1042)
JOURNAL     NIH-MGC http://mgc.ncl.nih.gov/.
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
             Unpublished (1999)
             Contact: Robert Strausberg, Ph.D.
             Email: cgabbs-remail.nih.gov
             Tissue Procurement: DCTD/DTP
             CDNA Library Preparation: Ling Hong/Rubin Laboratory
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
             Plate: LLCM518 row: b column: 07
             High quality sequence stop: 667.
             Location/Qualifiers
               1..1042
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone_image="3835686"
                 /clone_lib="NIH-MGC_9"
                 /tissue_type="adenocarcinoma cell line"
                 /lab_host="DH10B (phage-resistant)"
                 /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                 EcoRI; CDNA made by oligo-dT priming. Directionally
                 cloned into EcoRI/XhoI sites using the following 5'
                 adaptor: GGCAAGAG(G). Size-selected >500bp for average
                 insert size 1.8kb. Library constructed by Ling Hong in
                 the laboratory of Gerald M. Rubin (University of
                 California, Berkeley) using ZAP-cDNA synthesis kit
                 (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT      371 a 159 c 390 g 122 t
ORIGIN
```

/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 118 a 57 c 131 g 39 t

ORIGIN

Query Match 100.0%; Score 30; DB 11; Length 345;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
|||||

DB 13 GACATCAGCAGCAGCAGCAGCAGCAGC 42

RESULT 2
BF156008 516 bp mRNA EST 30-OCT-2000
LOCUS BF156008
DEFINITION RC0-HT0955-270900-033-d08 HT0955 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF156008
VERSION BF156008.1 GI:11051191
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 516)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=4t2-RC0-HT0955-270900-033-d08&ts=2000-09-27&ft=1)

Seq primer: puc 18 forward

High quality sequence start: 56

High quality sequence stop: 516.

FEATURES

source

1. 516

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HT0955"

/dev_stage="Adult"

/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

155 a 109 c 175 g 77 t

BASE COUNT

ORIGIN

Query Match 100.0%; Score 30; DB 11; Length 516;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
|||||

DB 245 GACATCAGCAGCAGCAGCAGCAGCAGC 274

RESULT 3
BG822824 662 bp mRNA EST 22-MAY-2001
LOCUS BG822824
DEFINITION 60272781661 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4867443 5',
ACCESSION BG822824
VERSION BG822824.1 GI:14170411
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 662)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L10M1735 row: P column: 04

High quality sequence stop: 601.
Location/Qualifiers

FEATURES

source

1. 662

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4867443"

/clone_lib="NIH_MGC_15"

/tissue_type="adenocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 202 a 139 c 255 g 66 t

ORIGIN

Query Match 100.0%; Score 30; DB 11; Length 662;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
|||||

DB 253 GACATCAGCAGCAGCAGCAGCAGCAGC 282

RESULT 4
BE742222 723 bp mRNA EST 15-SEP-2000
LOCUS BE742222
DEFINITION 601575629F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3836647 5',
RNA sequence.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:14:18 ; Search time 3881.49 Seconds
(without alignments)
83.054 Million cell updates/sec

Title: US-09-915-060-4

Perfect score: 30

Sequence: 1 gacatcagcagacagcagagagagacacagc 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
EST : *
1: em_estfun : *
2: em_esthum : *
3: em_estin : *
4: em_estom : *
5: em_estpl : *
6: em_estba : *
7: em_estro : *
8: em_estov : *
9: em_hic : *
10: qb_estl : *
11: qb_est2 : *
12: qb_hic : *
13: qb_gss : *
14: em_gss_fun : *
15: em_gss_hum : *
16: em_gss_inv : *
17: em_gss_pln : *
18: em_gss_pro : *
19: em_gss_rtd : *
20: em_gss_vrt : *
21: em_gss_other : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	345	11	BI261348 602953350
2	30	100.0	516	11	BF156008 RC0-HT095
3	30	100.0	662	11	BG822824 602727816
4	30	100.0	723	10	BE742222 601575629
5	30	100.0	740	10	AL043340 DKEP24340
6	30	100.0	1042	10	BE742943 601574925
7	30	100.0	1063	10	BE740559 601595656
8	30	100.0	1340	11	BG107702 602277878
9	26.8	89.3	182	10	AA413247 v652612.1
10	26.8	89.3	372	10	AL034747 v8608b53
11	26.8	89.3	649	10	AW210026 u150a11.1
12	26.8	89.3	900	11	BG298479 602396747

c	13	26.8	89.3	972	13	CNS0431P	AL272842 Tetradon
	14	26.8	89.3	1000	12	AK008283	AK008283 Mus muscu
	15	25.8	86.0	216	10	AV105652	AV105652 AV105652
	16	25.8	86.0	251	10	AV270235	AV270235 AV270235
	17	25.8	86.0	253	10	AV327502	AV327502 AV327502
	18	25.2	84.0	233	10	BB290670	BB290670 BB290670
	19	25.2	84.0	367	10	AA681955	AA681955 v144h10.s
	20	25.2	84.0	493	11	B1018825	B1018825 IL3-MT026
	21	25.2	84.0	758	11	BG669359	BG669359 602836841
	22	25.2	84.0	812	11	BI149024	BI149024 602911290
	23	25.2	84.0	830	11	BF607177	BF607177 M1_00011
	24	24.2	80.7	147	10	AV065528	AV065528 AV065528
	25	24.2	80.7	192	10	AV368285	AV368285 AV368285
	26	22.6	75.3	645	13	CNS01FLJ	AL141992 Anopheles
	27	22.6	75.3	651	13	CNS01K2L	AL147806 Anopheles
	28	22.6	75.3	681	13	CNS01G8W	AL143553 Anopheles
	29	22.6	75.3	685	13	CNS01M6Q	AL150547 Anopheles
	30	22.6	75.3	721	13	CNS01F9Q	AL141567 Anopheles
	31	22.6	75.3	750	13	CNS01FRW	AL142221 Anopheles
	32	22.6	75.3	762	10	AJ281330	AJ281330 4A3A-P2A8
	33	22.6	75.3	788	13	CNS01EBU	AL140239 Anopheles
	34	22.6	75.3	1011	13	CNS01FEW	AL143581 Anopheles
	35	22.6	75.3	1024	13	CNS01FE6	AL141461 Anopheles
	36	22.6	75.3	1151	11	BF140125	BF140125 601786074
	37	22	73.3	566	13	AQ400706	AQ400706 HS_5069_A
	38	21.2	70.7	985	13	CNS05T01	AL352522 Tetradon
	39	21.2	70.7	1050	13	CNS04UP5	AL308066 Tetradon
	40	21	70.0	415	10	AI515282	AI515282 LD47038.5
	41	21	70.0	513	10	AA390276	AA390276 LD08974.5
	42	21	70.0	547	13	AQ012093	AQ012093 575P1H1.07
	43	21	70.0	570	10	AA951266	AA951266 LD31649.5
	44	21	70.0	571	10	AA979394	AA979394 LD33942.5
	45	21	70.0	586	11	BG985824	BG985824 515513 NI

ALIGNMENTS

RESULT 1
BI261348
LOCUS 345 bp mRNA
DEFINITION 602953350F1 NIH_MGC_99 Homo sapiens CDNA clone IMAGE:5087556 5',
B1261348
ACCESSION BI261348
VERSION BI261348.1 GI:14820527
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Louis Staudt, M.D., Ph.D. CDNA library
Preparation: Ling Hong/Rubin laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1845 row: c column: 13
High quality sequence stop: 345.
Location/Qualifiers

FEATURES

source
1.345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5087556"
/clone_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"


```
BASE COUNT      666 a      576 c      741 g      357 t
ORIGIN
      SEQORVKRGTSPRPREGGLGYSQGGDDDKETGPHLTTTNOGASAAAGPFSILKF"

Query Match      100.0%; Score 30; DB 9; Length 2340;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 gacatcagcagcagcaggaagaccagc 30
      |||
      776 GACATCAGCAGCAGCAGGAGAGACCAGC 805

RESULT 14
AF067517      2349 bp      mRNA      PRI      07-NOV-1998
LOCUS      Homo sapiens PITSURE protein kinase alpha SV11 isoform (CDC2L1)
DEFINITION      mRNA, complete cds.
ACCESSION      AF067517
VERSION      AF067517.1 GI:3850313
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 2349)
      Gurrarajan, R., Lahli, J.M., Grenet, J., Easton, J., Gruber, I.,
      Ambros, P.F. and Kidd, V.J.
      Duplication of a genomic region containing the Cdc2L1-2 and
      MPM21-22 genes on human chromosome 1p36.3 and their linkage to D122
      Genome Res. 8 (9), 929-939 (1998)
      98424414
      2 (bases 1 to 2349)
      Gurrarajan, R., Lahli, J.M., Grenet, J., Easton, J., Gruber, I.,
      Ambros, P. and Kidd, V.J.
      Direct Submission
      Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
      Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
      Location/Qualifiers
      source
      1. .2349
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="1"
      /map="1p36.3"
      1. .2349
      /gene="CDC2L1"
      626. .2323
      /gene="CDC2L1"
      /note="corresponding genomic sequence deposited as GenBank
      Accession Numbers AF080678-AF080688, AF092429, and
      AF092430"
      /codon_start=1
      /product="PITSURE protein kinase alpha SV11 isoform"
      /protein_id="AAC72082.1"
      /db_xref="GI:3850314"
      /translation="MREYSDKVKASHWSRPPRRRPFELGGRKPKVEKKEEERD
      LLSLDLSDISDERKTSASSESSAEGSSGSEEEEEEESTSESESEEEEEE
      EEEETGNSSESESESESESESESESESESESESESESESESESESESE
      VEGTPOSSALTEGDDYVDPALSPILKQELPKYLPALQCSVEEEOCINRIEEGT
      YGVYRARDKKTDEIVALKRLEKMEKEGEPITSLREINTILIKOHPNIVYRILVYG
      SMMDKITVYNNVYEHDLKSLMETKOPFLPEGVKTIQLRGVXKHLHNDNITLRDLK
      TSNLLSHAGILTKVGDGLARVGSPLKATPVVVTWYRAPELILGAKETSTAVDMV
      SVGCIIFGELLTKPLPFGKSEIDQINKVFKDGLTPESEKIMPGSELPAVKKMTSEHP
      YNNLRFGALLSDQFDLNNKFLTYFPGRRISAEGLKHEYFRFTPLPIDPSMPTW
      PAKSEORVKRGTSPRPREGGLGYSQGGDDDKETGPHLTTTNOGASAAAGPFSILKF"

BASE COUNT      668 a      577 c      748 g      356 t
ORIGIN

Query Match      100.0%; Score 30; DB 9; Length 2349;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 gacatcagcagcagcaggaagaccagc 30
      |||
      776 GACATCAGCAGCAGCAGGAGAGACCAGC 805

RESULT 15
AF067523      2362 bp      mRNA      PRI      07-NOV-1998
LOCUS      Homo sapiens PITSURE protein kinase beta SV7 isoform (CDC2L2) mRNA,
DEFINITION      complete cds.
ACCESSION      AF067523
VERSION      AF067523.1 GI:3850325
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 2362)
      Gurrarajan, R., Lahli, J.M., Grenet, J., Easton, J., Gruber, I.,
      Ambros, P.F. and Kidd, V.J.
      Duplication of a genomic region containing the Cdc2L1-2 and
      MPM21-22 genes on human chromosome 1p36.3 and their linkage to D122
      Genome Res. 8 (9), 929-939 (1998)
      98424414
      2 (bases 1 to 2362)
      Gurrarajan, R., Lahli, J.M., Grenet, J., Easton, J., Gruber, I.,
      Ambros, P. and Kidd, V.J.
      Direct Submission
      Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
      Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
      Location/Qualifiers
      source
      1. .2362
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="1"
      /map="1p36.3"
      1. .2362
      /gene="CDC2L2"
      1143. .2336
      /gene="CDC2L2"
      /note="corresponding genomic sequence deposited as GenBank
      Accession Numbers AF080689-AF080697 and AF092426-AF092428"
      /codon_start=1
      /product="PITSURE protein kinase beta SV7 isoform"
      /protein_id="AAC72088.1"
      /db_xref="GI:3850326"
      /translation="MKNEKKTSMFLFOSHSSTEIRGVYKORRKKMGSRVEEFOCLN
      KIEGTVGYVYRAKDKKTDEIVALKRLEKMEKEGEPITSLREINTILIKOHPNIVY
      RETVGSNMDKITVYNNVYEHDLKSLMETKOPFLPEGVKTIQLRGVXKHLHNDNIT
      LRDLKTSNLLSHAGILTKVGDGLARVGSPLKATPVVVTWYRAPELILGAKETSTAVDMV
      TAVDMVSVGCIIFGELLTKPLPFGKSEIDQINKVFKDGLTPESEKIMPGSELPAVKK
      TFESEHYNNLRKRFEGALLSDQFDLNNKFLTYFPGRRISAEGLKHEYFRFTPLPIDP
      SMPTWPAKSEORVKRGTSPRPREGGLGYSQGGDDDKETGPHLTTTNOGASAAAGP
      FSILKF"

BASE COUNT      681 a      562 c      770 g      349 t
ORIGIN

Query Match      100.0%; Score 30; DB 9; Length 2362;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Search completed: December 4, 2001, 18:54:31
Job time: 9167 sec

```
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2323
/gene="CDC2L2"
1104..2297
/feature="CDC2L2"
/feature="corresponding genomic sequence deposited as GenBank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PITSLRE protein kinase beta SV8 isoform"
/protein_id="AAC72083.1"
/db_xref="GI:3850316"
/translation="MRNEKMTTSMLFQSHSTFEPGRVKQKQKRWKGRSVEEFOCLN
RIEETGYGVVYRAKDKDEIYALKRIMKEKEGEPITSLREINTILKAQHNTVY
REIVGSSNMKITYVMNYVHDLKSLMEKQEPILGKVTLMIOILRGVKKLHDMKI
LHRDLKTSNLLSHAGILKVGDLAREGSPKATTPVYVITWYTAPELLIGAKYS
TAYDMSVGCITFGEILFQKPLFPNGSEIDINKVRELSTPSKRWGSELPVAKYS
TFSEHPYNNLRKREGALLSDQGFDMNKFLTYPPGRRISAEDGLKHEYRETPIDP
SMPTWPAKSEQDVRKRGTSPPREGGLGYSQGLDDELKETGHLTTYNQGSAAAPGC
FSLKF"

BASE COUNT      673 a      550 c      754 g      346 t
ORIGIN

Query Match      100.0%; Score 30; DB 9; Length 2323;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 gacatcagcagcagcagcagcagcagc 30
      ||||||||||||||||||||||||||||
Db      875 GACATCAGCGACGACGAGAGAGACCACG 904

RESULT 12
AF067525      2329 bp      mRNA      PRI      07-NOV-1998
LOCUS      Homo sapiens PITSLRE protein kinase beta SV13 isoform (CDC2L2)
DEFINITION      mRNA, complete cds.
ACCESSION      AF067525
VERSION      AF067525.1 GI:3850329
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 2329)
AUTHORS      Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE      Duplication of a genomic region containing the Cdc2L1-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to D122
JOURNAL      Genome Res. 8 (9), 929-939 (1998)
FEATURES
source
1..2329
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2329
/gene="CDC2L2"
68..1654
/feature="CDC2L2"
/feature="corresponding genomic sequence deposited as GenBank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PITSLRE protein kinase beta SV13 isoform"
```

```
/protein_id="AAC72090.1"
/db_xref="GI:3850330"
/translation="MEITIRNSPYRREDSMEDRGEEDSLAIKPPQKSNKEKQVHNRK
DEKREKMKHARVKEHEHRRKRRNEEDOKAREREROKRREARERARERARERAGNDGFC
LEFRDLQLELERRERERERERERERERERERERERERERERERERERERERERERERER
DYSDVKVASHMSRSPRPREFELGDKRKVKEKMEERDLSDLODSDSEKTS
AESSSAESGSGSEEEEEESEEGSTSESEEESEETGSSNESAESAEEV
SEESSEDEBERNENHLLVPESSRDNRSGSEFAEEVGCSTPOSSALPEADVPDS
PALILTELKQELPKYLPALOCRSVEEQCLNRIEETGYVYTAOKDKDEIYALKR
IKMEKEKGFPTSLREINTILKAQHNTVYREIVGSSNMKITYVMNYVHDLKSL
METMKQPLPGEVKTLMQLRGVKKHLHDMNLIHRDLKTSNLLSHAGILKVSPPSG
PSQDPPPTPSRPSVAG"

BASE COUNT      634 a      605 c      738 g      352 t
ORIGIN

Query Match      100.0%; Score 30; DB 9; Length 2329;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 gacatcagcagcagcagcagcagcagc 30
      ||||||||||||||||||||||||||||
Db      689 GACATCAGCGACGACGAGAGACCACG 718

RESULT 13
HSU04818      2340 bp      mRNA      PRI      08-JUL-1994
LOCUS      Human protein kinase PITSLRE alpha 2-4 mRNA, complete cds.
DEFINITION      U04818
ACCESSION      U04818
VERSION      U04818.1 GI:507163
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 2340)
AUTHORS      Xiang,J., Lahli,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE      Molecular cloning and expression of alternatively spliced PITSLRE
protein kinase isoforms
JOURNAL      J. Biol. Chem. 269, 15786-15794 (1994)
MEDLINE      94253170
REFERENCE      2 (bases 1 to 2340)
AUTHORS      Kidd,V.J.
TITLE      Direct Submission
JOURNAL      Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's
Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
Memphis, TN 38101, USA
FEATURES
source
1..2340
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1p36-2"
/cell_line="Hela S3"
/cell_type="epitheloid"
/tissue_type="cervix"
/clone_lib="Hela CDNA library"
/seq_stage="adult"
1..2340
626..2314
/codon_start=1
/product="PITSLRE alpha 2-4"
/protein_id="AAA19584.1"
/db_xref="GI:507164"
/translation="MRDYSKVAASHMSRSPRPREFELGDKRKVKEEKEERD
LLSDLODSDSEKTSAESSSAESGSGSEEEEEESEEGSTSESEEESEEE
ETGSSNESAESAEEVSEESSEDEBERNENHLLVPESSRDNRSGSEFA
GTPOSSALPEADVPDSALPEADVPDSALPEADVPDSALPEADVPDSALPEADVPDS
VYRAKDKDEIYALKRIMKEKEGEPITSLREINTILKAQHNTVYREIVGSSNM
DKITYVMNYVHDLKSLMEKQEPILGKVTLMIOILRGVKKHLHDMNLIHRDLKTSN
LLSHAGILKVGDLAREGSPKATTPVYVITWYTAPELLIGAKYSTAVDMWSVG
CIFGELLFQKPLFPNGSEIDINKVRELSTPSKRWGSELPVAKYSSTAVDMWSVG
LRRKREGALLSDQGFDMNKFLTYPPGRRISAEDGLKHEYRETPIDPMPPTWPAK
```

AUTHORS	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J.
TITLE	Direct Submision
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES	location/Qualifiers
SOURCE	1..2226
gene	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="1" /map="1p36.3"
CDS	1..2226 /gene="CDC2L1" 620..2200 /gene="CDC2L1" /note="corresponding genomic sequence deposited as Genbank Accession Numbers AF080678-Af080688, AF092429, and AF092430" /codon_start=1 /product="PITSLR protein kinase alpha SV4 isoform" /protein_id="AAC72078.1" /db_xref="GI:3850306" /translation="MEERDLSDLDIDISDERKTSAAESSAASGSGSEEEEEESEEEEESEESSEEESEETGSSEASQSAEVSSEEMSEEDERNENLLMLPKYLPALVESNFDPSDSEAEFEVEEGTPOSSALTEDVDYDPALSPTEIKOELPKYLPALOCRSVEFDCGLNRLENGTVGYRRADKATDEIVALKRLMKEREKEGPITSLREITTLRKOHNVITVREIIVYGSMOKITYVMNTVSHDKLSMETMROPFUPEGVKTLMTILRGKHLHDHWILLRDLKTSNLISHAGILTKVGDGFLAREGSPKATPVVYLWLMRAPPELLAKKEYSTAVDMWSVGCIFGELLTORPLFGKSEIDQINKVFEDLTGPEKMGVESELPAVKRMFTSEHPYNLKRFGALLSDQGDLNNKFETYFPGRRIISAEGDHLHYEFEPFLPIDPSMFPTPMWPAKSPHQQRVKNRTSRPRPEGGLGYSQIGDDDLKETGFHTTTNGASACAPGFSIMK"
BASE COUNT	638 a 541 c 705 g 342 t
ORIGIN	
Query Match	100.0%; Score 30; DB 9; Length 2226;
Best Local Similarity	100.0%; Pred. No. 0.28;
Matches	30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 gacatcagcgacagcgagaaggaccgcg 30 Db 653 GACATCAGCGACAGCGAGAAGACCAGC 682
RESULT 11	
AF067518	
LOCUS	AF067518 2323 bp mRNA PFI 07-NOV-1998
DEFINITION	Homo sapiens PITSLR protein kinase beta sv8 isoform (CDC2L2) mRNA,
ACCESSION	Complete cds.
VERSION	AF067518
KEYWORDS	AF067518.1 GI:3850315
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 2323)
AUTHORS	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P.F. and Kidd,V.J.
TITLE	Duplication of a genomic region containing the Cdc2L1-2 and
JOURNAL	MMP21-22 genes on human chromosome 1p36.3 and their linkage to D12Z2
MEDLINE	Genome Res. 8 (9), 929-939 (1998) 98424414
REFERENCE	2 (bases 1 to 2323)
AUTHORS	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J.
TITLE	Direct Submission
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES	location/Qualifiers
SOURCE	1..2323
ORGANISM	"Homo sapiens"

Query Match 100.0%; Score 30; DB 6; Length 468;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcagcgacgaggaagaccagc 30
|||||
Db 439 GACATCAGCGACGAGAGGAGACACGAC 468

RESULT 6
LOCUS CDC2L2S08 480 bp DNA PRI 09-DEC-1998
DEFINITION Homo sapiens PITSURE protein kinase (CDC2L2) gene, exons 8 and 9.
ACCESSION AF080678
VERSION AF080678.1 GI:3978432
KEYWORDS
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 480)
AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to DL22
JOURNAL Genome Res. 8 (9), 929-939 (1998)
MEDLINE 98424414
REFERENCE 2 (bases 1 to 480)
AUTHORS Gururajan,R., Lahti,J.L., Grenet,J., Easton,G., Gruber,I.,
Ambros,P. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-1998) Tumor Cell Biology, St Jude Children's
Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
LOCATION/Qualifiers
FEATURES
source 1..480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
78..173
/gene="CDC2L1"
/number=8
259..417
exon /gene="CDC2L1"
/number=9
BASE COUNT 136 a 99 c 156 g 89 t
ORIGIN

Query Match 100.0%; Score 30; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. NO. 0.36;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcagcgacgaggaagaccagc 30
|||||
Db 125 GACATCAGCGACGAGAGGAGACGACG 154

RESULT 7
LOCUS CDC2L2S08 592 bp DNA PRI 13-DEC-1998
DEFINITION Homo sapiens protein kinase PITSURE (CDC2L2) gene, exons 8 and 9.
ACCESSION AF080689
VERSION AF080689.1 GI:4007426
KEYWORDS
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 592)
AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,

Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to DL22
JOURNAL Genome Res. 8 (9), 929-939 (1998)
MEDLINE 98424414
REFERENCE 2 (bases 1 to 592)
AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-1998) Tumor Cell Biology, St. Jude Children's
Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
LOCATION/Qualifiers
FEATURES
source 1..592
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
78..173
/gene="CDC2L2"
/number=8
259..408
exon /gene="CDC2L2"
/number=9
BASE COUNT 154 a 126 c 184 g 128 t
ORIGIN

Query Match 100.0%; Score 30; DB 9; Length 592;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcagcgacgaggaagaccagc 30
|||||
Db 125 GACATCAGCGACGAGAGGAGACGACG 154

RESULT 8
LOCUS AX033425 660 bp DNA PAT 21-SEP-2000
DEFINITION Sequence 6 from Patent WO0044896.
ACCESSION AX033425
VERSION AX033425.1 GI:10280186
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Beyaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
JOURNAL Patent: WO 0044896-A 6 03-ANG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE); BEYAERT RUDI (BE); CORNELIS
SIGRID (BE)
LOCATION/Qualifiers
FEATURES
source 1..660
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 210 a 128 c 261 g 61 t
ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 660;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcagcgacgaggaagaccagc 30
|||||
Db 439 GACATCAGCGACGAGAGGAGACGACG 468

RESULT 9
AF174497

BASE COUNT 11 a 8 c 10 g 1 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 30; DB 6; Length 30;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
1 GACATCAGCGACGACGAGAGAGACCAGC 30

RESULT 2
AX033439 40 bp DNA PAT 21-SEP-2000
LOCUS AX033439
DEFINITION Sequence 20 from Patent W00044896.
ACCESSION AX033439
VERSION AX033439.1 GI:10280200
KEYWORDS
SOURCE
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 40)
AUTHORS Beyaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and uses thereof
JOURNAL Patent: WO 0044896-A 20 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS SIGRID (BE)

FEATURES
source Location/Qualifiers
1..40
/organism="Homo sapiens"
/db_xref="taxon:32630"
/note="5'-end primer"

BASE COUNT 14 a 10 c 12 g 4 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 30; DB 6; Length 40;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
11 GACATCAGCGACGACGAGAGAGACCAGC 40

RESULT 3
AX033420 222 bp DNA PAT 21-SEP-2000
LOCUS AX033420
DEFINITION Sequence 1 from Patent W00044896.
ACCESSION AX033420
VERSION AX033420.1 GI:10280181
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 222)
AUTHORS Beyaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and uses thereof
JOURNAL Patent: WO 0044896-A 1 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS SIGRID (BE)

FEATURES
source Location/Qualifiers
1..222
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 77 a 36 c 89 g 20 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 30; DB 6; Length 222;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
1 GACATCAGCGACGACGAGAGAGACCAGC 30

RESULT 4
AX033421 222 bp mRNA PAT 21-SEP-2000
LOCUS AX033421
DEFINITION Sequence 2 from Patent W00044896.
ACCESSION AX033421
VERSION AX033421.1 GI:10280182
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 222)
AUTHORS Beyaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and uses thereof
JOURNAL Patent: WO 0044896-A 2 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS SIGRID (BE)

FEATURES
source Location/Qualifiers
1..222
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 77 a 36 c 89 g 20 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 30; DB 6; Length 222;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
1 GACATCAGCGACGACGAGAGAGACCAGC 30

RESULT 5
AX033424 468 bp DNA PAT 21-SEP-2000
LOCUS AX033424
DEFINITION Sequence 5 from Patent W00044896.
ACCESSION AX033424
VERSION AX033424.1 GI:10280185
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 468)
AUTHORS Beyaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and uses thereof
JOURNAL Patent: WO 0044896-A 5 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS SIGRID (BE)

FEATURES
source Location/Qualifiers
1..468
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 144 a 100 c 182 g 42 t
ORIGIN

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:54:30 ; Search time 2399.05 Seconds
(without alignments)
206.296 Million cell updates/sec

Title: US-09-915-060-4

Sequence: 1 gacatcagcgacagcgagaggaagaccagc 30

Scoring table: IDENTITY_NUC

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Listing first 45 summaries

Database :

```

1:  gb.ba:*
2:  gb.htg:*
3:  gb.in:*
4:  gb.om:*
5:  gb.ov:*
6:  gb.pat:*
7:  gb.ph:*
8:  gb.pl:*
9:  gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.com:*
20: em.or:*
21: em.ov:*
22: em.pat:*
23: em.ph:*
24: em.pl:*
25: em.ro:*
26: em.sts:*
27: em.sy:*
28: em.un:*
29: em.vi:*
30: em.hqo.hum:*
31: em.hqo.luv:*
32: em.hqo.rod:*
33: em.hqo.hum:*
34: em.hqo.luv:*
35: em.hqo.rod:*
36: em.hqo.other:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	30	100.0	30	6	AX033423	AX033423 Sequence
2	30	100.0	40	6	AX033439	AX033439 Sequence
3	30	100.0	222	6	AX033440	AX033440 Sequence
4	30	100.0	222	6	AX033421	AX033421 Sequence
5	30	100.0	468	6	AX033424	AX033424 Sequence
6	30	100.0	480	9	CD021308	AF080678 Homo sapi
7	30	100.0	552	9	CD021250	AF080688 Homo sapi
8	30	100.0	660	6	AX033425	AF080678 Homo sapi
9	30	100.0	1715	9	AF174497	AF080688 Homo sapi
10	30	100.0	2226	9	AF067513	AF174497 Homo sapi
11	30	100.0	2323	9	AF067518	AF067513 Homo sapi
12	30	100.0	2329	9	AF067525	AF067518 Homo sapi
13	30	100.0	2340	9	HS004818	AF067525 Homo sapi
14	30	100.0	2349	9	HS004818	AF067525 Homo sapi
15	30	100.0	2362	9	AF067523	U04818 Human prote
16	30	100.0	2409	9	AF067521	AF067517 Homo sapi
17	30	100.0	2436	9	AF067520	AF067521 Homo sapi
18	30	100.0	2439	9	AF067519	AF067521 Homo sapi
19	30	100.0	2444	9	HS004817	AF067520 Homo sapi
20	30	100.0	2448	9	AF067522	U04817 Human prote
21	30	100.0	2465	9	HS007705	AF067522 Homo sapi
22	30	100.0	2471	6	AX033422	U07705 Human prote
23	30	100.0	2471	9	HS004816	AX033422 Sequence
24	30	100.0	2471	9	HS007704	U04816 Human prote
25	30	100.0	2477	9	HS004824	U07704 Human prote
26	30	100.0	2486	9	AF067512	U04824 Human prote
27	30	100.0	2500	9	AF067514	AF067512 Homo sapi
28	30	100.0	2525	9	AF067515	AF067512 Homo sapi
29	30	100.0	2533	9	AF067516	AF067515 Homo sapi
30	30	100.0	2544	9	AK000081	AF067516 Homo sapi
31	30	100.0	110608	9	HS283823	AK000081 Homo sapi
32	28	89.3	35	6	AX033437	AL031282 Human DNA
33	26.8	89.3	1936	6	M0SP55GTA	AX033437 Sequence
34	26.8	88.3	3161	10	MUSCDK	M56633 Mouse p58/
35	23.6	78.7	1670	10	RATGTA	L37092 Mus musculi
36	22.6	75.3	8071	1	MS00T19T	L24398 Rattus nor
37	22.6	75.3	329709	1	AP002297	M93690 Mesocricet
38	22.2	73.3	248	5	CP17SLE04	AP002297 Mesocricet
39	22.2	73.3	82852	2	AC017665	U16347 Gallus galli
40	22.2	73.3	172414	2	AC093168	AC017665 Drosophil
41	22.2	73.3	199607	2	AC009366	AC093168 Homo sap
42	22.2	73.3	295566	3	AE003597	AC009366 Drosophil
43	21.8	72.7	151589	8	AC079022	AE003597 Drosophil
44	21.2	70.7	217594	8	AP002066	AC079022 Oryza sat
45	21	70.0	4281	3	AY038001	AP002066 Agrobacte
						AY038001 Drosophil

ALIGNMENTS

RESULT 1	
AX033423	
LOCUS	AX033423 30 bp DNA
DEFINITION	Sequence 4 from Patent WO0044896.
ACCESSION	AX033423
VERSION	AX033423.1 GI:10280184

ORGANISM

REFERENCE
1 (bases 1 to 30)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE

uses thereof.

JOURNAL Patent: WO 0044896-A 4 03-AUG-2000;

FEATURES	Location/qualifiers
source	1. .30

Trace considered overall poor quality
High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers

1. .367

/organism="Mus musculus"

/strain="B6D2 F1/J"

/db_xref="taxon:10090"

/clone="IMAGE:1123555"

/clone_lib="Knowles Solter mouse 2 cell"

/tissue_type="embryo"

/dev_stage="2-cell"

/lab_host="DH10B"

/note="Organ: embryo; Vector: pBluescribe (modified);

Site_1: MluI; Site_2: SalI; Cloned unidirectionally from

mRNA prepared from 13,500 2-cell stage embryos. Primer:

SalI(dT): 5'-CGGTGACCGCTGACCGCTTTTCTTTT-3'. CDNAS

were cloned into the MluI/SalI sites of a modified

pBluescribe vector using commercial linkers (NEB).

Average insert size: 1.2 kb."

BASE COUNT

111 a

68 c

143 g

45 t

ORIGIN

Query Match

39.9%; Score 186.6; DB 10; Length 367;

Best Local Similarity 76.3%; Pred. No. 3.1e-24;

Matches 242; Conservative 0; Mismatches 74; Indels 1; Gaps 1;

```
OY 153 aaggaagcggagcggagcggagcgaatgaggagcagcagaagagcggagcagcagaagcaga 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1  AAGGAAGAGGAGCGGAGAAAGACGTGACGGAGCAGCAGAAAGAGCAGCGGAGCAACAA 60

OY 213 ggaagcggagcggagcggagcggagcgaagagcggagcggagcggagcggagcggagc 272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 GGAGCGGGAACGAGGCGCAGAGGGCGCCACAAAGAGAGAGAACGCCGTAGGAGGTGTC 120

OY 273 tgcacatcccaagcgaatgagagagactacagcgacaaagtgaagccagccactgag 332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 TCGGCAATCCGCTACCGTGAAGGAGAGTGTGATGATGGAAGTTGGCCACTGGA 180

OY 333 tcgcaagccgcctcgccgcgc-gcggaagcgttcgaglttggaagcgcgcgaagccag 391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 CCGCATCCGTCGGAGGCCACCGAGAGAGCGCTGTCAAGTGGAGGCAACCGAGAGCCAG 240

OY 392 taaaagaagaagaatggaagaaggaactgtgtccgaacttacagagacatcagcgaca 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 TAAAGAAAGAGAGTGAAGAGAGAGAGCTGTGTCAAGCTCCAGACATCACTAGTGACA 300

OY 452 gcgagaggaagaccagc 468
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 GCGAGAGGAAGAGAGC 317
```

Search completed: December 4, 2001, 18:14:23
Job time: 6764 sec


```

/clone="IMAGE:4394166"
/clone_1id="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: pCMV-SPORE6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; 0190-df primed. Average insert size 1,383 kb. Library enriched for full-length clones and constructed by Life Technologies Note: this is a NIH_MGC library."
BASE COUNT      297 a      154 c      272 g      103 t
ORIGIN

```

Query Match	56.8%;	Score 265.8;	DB 11;	Length 826;
Best Local Similarity	85.9%;	Pred. No. 2.2e-38;		
Matches 353;	Conservative	0;	Mismatches 27;	Indels 31;
				Gaps 4

QY	1	cacgaacgctcggaaacacatcggaaagaaacagatataaactcgcgcggatbtgaaaga	60
Db	430	CACGAACGTCGGAACGGCATCGAAGAAACAGGATTAACCTCGCCGGAAATGGGAAGA	489
QY	61	cagaagaagagggaaatcgcaagggagcatctcagaagagaaggggaaatgatgycgtg	120
Db	490	CAGAAGAAGAGGGGAGTGGCAGAGGACATCTCCAGGAGAGAA-----	531
QY	121	tgctcttcagagagccgtcttgaaacagttaagaagagagcgcgagcg-ccaagt	179
Db	532	-----AGGACCCCTTGGACCACTTATGAAAGCAAGCGGAGCGCGGCCCAAGAT	582
QY	180	gcggagacagcagaagaagacagcggagacagaagaaacgagccgcg-99gcggagagc	238
Db	583	GCGGGACACACAGAAGAGCAGCAGCGGAGCAGAAAGAGACCGGACGGCGGGCCGAGGAGC	642
QY	239	gcgcgaagagcgcgagggcccgcaagggaaagtgtctgcacataccccaagcatagagag	298
Db	643	GGCCCAAGGAGCGGAGGCCCGCAGGGAATGTCGACATCAACGAAACATAGAGAGG	702
QY	299	actaagcgcgacaagtgaaagccagccactcgaggtcgacgccgcctcgccgcgcgcgg	358
Db	703	ACTAAGCAGCAAAAGTAAAAACCAAGCCACCTGAGATCGCACCCGCCCTCGGC--GCCGCG	760
QY	359	acgcgttcaggtctggagacagcgccggaagccagttaaagagagaataatgg	409
Db	761	GACGCTTCGAGTTGGGAAGCCCGGACCACTAAGAAAGAAAAAAGGGG	811

RESULT	12
LOCUS	BE259507
DEFINITION	BE259507 674 bp mRNA EST 13-JUL-2000
ACCESSION	G01106336F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3342948 5',
VERSION	BE259507
KEYWORDS	mRNA sequence.
SOURCE	BE259507.1 GI:9129943
ORGANISM	EST.
JOURNAL	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
COMMENT	1 (bases 1 to 674)
	NIH-MGC http://mgc.nci.nih.gov/ .
	National Institutes of Health, Mammalian Gene Collection (MGC)
	unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: cgabbs-r@mail.nih.gov
	Tissue Procurement: ATCC
	cDNA library Preparation: Ling Hong/Rubin Laboratory
	cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA Sequencing by: Incyte Genomics, Inc.
	Clone distribution: MGC clone distribution information can be
	found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
	Plate: LLCM126 row: o column: 13
	High quality sequence start: 37
	High quality sequence stop: 672.

FEATURES	Location/Qualifiers
source	1. .674

```

/lab-host="Ddl10B (phage-resistant)"
/notice="Organ: eye; Vector: pOT81; Site 1: XhoI; site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: This is a NIH-MGC library."

```

BASE COUNT	248 a	115 c	229 g	82 t
ORIGIN				

Query Match	43.9%;	Score 205.6;	DB 10;	Length 674;
Best Local Similarity	88.9%;	Pred. No. 1.2e-27;		
Matches 256;	Conservative	0;	Mismatches 4;	Indels 28;
				Gaps 2

[illegible]

RESULT	13
AL034747	
LOCUS	
DEFINITION	AL034747 372 bp mRNA EST 29-DEC-1999
ACCESSION	V8608B53 Beddington mouse dissected endoderm Mus musculus cDNA
VERSION	Clone 528_8K10 5', mRNA sequence.
KEYWORDS	AL034747 AL034747.1 GI:6646373
SOURCE	EST.
ORGANISM	house mouse. Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 372)
AUTHORS	Harrison,S.M., Dunwoodie,S.L., Arkell,R.M., Lehrach,H. and Beddington,R.S.
TITLE	Isolation of novel tissue-specific genes from cDNA libraries representing the individual tissue constituents of the gastrulating mouse embryo
JOURNAL	Development 121 (8), 2479-2489 (1995)
MEDLINE	93401865
COMMENT	Contact: Wiles,M., Lehrach,H. and Avner,P.

2, rue Gaston Cremlieux, 91000 Evry, France
Email: pavner@pasteur.fr
Clone available from Ressourcezentrum, Heubnerweg 6, D-14050
Berlin, Germany. Web site <http://www.rzpd.de>
Seq primer: CCGGTCCGGAAATTCCTCCGGT;

QY 61 cagaagagaaggaatgagcaaggagcattccagagagagaaggggaatgatgagcgtg 120
|||||
Db 418 CAGAAAGAGGAGGATGCGAAGGAGGACATTCCAGAGAGAA----- 459
QY 121 tgcctctcaggagccgcttgagcaagltcgaagaagcgagcgagcgagcgcaatg 180
|||||
Db 460 -----AGGAGCCCTTGAGCACTTAGAAAGAGCGGAGCGGAGCCGCAAGATG 510
QY 181 cggagagcagcaggaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 240
|||||
Db 511 CGGAGGAGCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 570
QY 241 cgcgaagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 300
|||||
Db 571 CGCAAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 630
QY 301 tacagcgacaaagtgaagccagcactgagtcgcagccgctcgagcgagcgagcg 360
|||||
Db 631 TACAGCGACAAAGTGAAAGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 690
QY 361 cggctcgaattggagagcg-cgaggaagcgagtgaaagaagagaagaatggaagaagga 419
|||||
Db 691 CGGTCGAGGTTGGAGAGCGGTGGGAGCGAGTAAAGAGCAGACACACTGAGCGAAGG 750
QY 420 cctgctgctccagactcaagagcatcagcgagcgagcgagcgagcgagcgagcg 458
|||||
Db 751 GGACCTCGCGGTCCGACTTTACAGGAGCATCAGGAGCAG 789

RESULT 7
Bg107702 1340 bp mRNA EST 30-JAN-2001
LOCUS 602217878F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4365474 5',
DEFINITION mRNA sequence.
ACCESSION Bg107702
VERSION Bg107702.1 GI:12601548
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1340)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@dsfemail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10014 row: 1 column: 19
High quality sequence stop: 667.
Location/Qualifiers
1. 1340
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4365474"
/clone_lib="NIH_MGC_86"
/tissue_type="osteocarcoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bone; Vector: PCMV-SPORE6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.533 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 467 a 272 c 450 g 151 t
ORIGIN

Query Match 65.2%; Score 305; DB 11; Length 1340;
Best Local Similarity 100.0%; Pred. No. 2.1e-45;
Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 agcgagagcgcaagatgctcgagcgagcagaagaagcgagcgagcgagcgagcgagcg 223
|||||
Db 1 AGCGGAGCGGAGGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
QY 224 ggcggcgaggaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 283
|||||
Db 61 GGGGGCGCGAGAGCG 120
QY 284 gaacgagtagagagagcctcagcgagcaagtgaaagcgagcgagcgagcgagcgagcg 343
|||||
Db 121 GAACGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 344 ctgcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 403
|||||
Db 181 CTGGCGCGCGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 404 aaatggaagaaggaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgag 463
|||||
Db 241 AATGGAAGAAAGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 464 ccagc 468
|||||
Db 301 CCAGC 305

RESULT 8
AK008283 1000 bp mRNA HTC 05-JUL-2001
LOCUS AK008283
DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length
enriched library, clone:2010016J05, full insert sequence.
ACCESSION AK008283
VERSION AK008283.1 GI:12842370
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male small intestine cDNA to
mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
clone:2010016J05.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1000)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (bases 1 to 1000)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,K., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome research. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (bases 1 to 1000)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Kono,H., Akiyama,U., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,
Sunu,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Itawa,K., Tanaka,T., Matsura,S., Kawai,D.,
Toneida,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,D.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipipillary sequencer
JOURNAL Genome research. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4 (bases 1 to 1000)
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.

[illegible]

RESULT	5
AI580044/c	
LOCUS	
DEFINITION	
AI580044	592 bp mRNA
	EST
	14-DEC-1999
tt445608 .x1 NCI CGAP Ut1 Homo sapiens cDNA clone IMAGE:2211806 3'	
similar to TR:061399 Q61399 CELL DIVISION CYCLE 2-LINE 2 ; contains	
element TARI repetitive element ; mRNA sequence.	

ACCESSION	AI580044	
VERSION	AI580044.1	GI:4564420
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>
1 (pages 1 to 592)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

JOURNAL • Tumor Gene Index
COMMENT • Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Mail: Bethesda, Maryland 20892

Email: cg405@remail.nlm.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1082 Std Error: 0.00
Seq Primer: -400P from Gibco
High quality sequence stop: 405
POLYA-NO.

FEATURES	SOURCE	
Location/Qualifiers		
1. .592		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/clone="IMAGE:2211806"		
/clone_lib="NCI_CGAP_Ut1"		
/tissue_type="well-differentiated endometrial adenocarcinoma, 7 pooled tumors"		
/lab_host="DH10b"		
/note="Organ: uterus; Vector: pCMV-Sport6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT Average insert size 1.75 kb. Life Technologies catalog # 11538-014"		
58 a	212 c	115 g 207 t
ORIGIN		

Query Match	71.18;	Score 332.8;	DB 10;	Length 592;
Best Local Similarity	92.88;	Pred. No. 2.3e-50;		
Matches 371; Conservative	0;	Mismatches 2;	Indels 27;	Gaps 1
QY 1 cccgaagctcggaacagcacatcgaagaacagataaagcttcgccgggaatggaaga	60			
DB 373 CACGACGCTCGGAACGGCATTGAGGAAGAACAGGATTAACCTTCGCCGGGAATGGGAAAA	314			
QY 61 cgaagaagaagggaatgcaagagagacatctcaagagagaagaagggaatgatgagctg	120			
DB 313 CAGGAAGAAAGGGAATGCGAAGGAGACATTCCAGGAGAAA-----	270			

QY	121	tgccctcttcagagacgcgccttgcagcaatttgaataagagagcgcgagcgcgagatg	180
Db	271	-----AAGGACCGCGTTGGACAGTTCAGAAAGGAAGCGGGAGCGCAAGATG	221
QY	181	cggagacgacgagaagagcagcgagagacaagaagacgcgagcgcgcgcgagagcgg	240
Db	220	CGGGCGAGCGCAAGAGAGACACCGGGACACGAAGAGGCGCGCGCGGGGAGACGG	161
QY	241	cgcgaagagcgagagagccgcgagagagtgtrctgcacataccgaacgatgtgagagagac	300
Db	160	CGCAAGAGACGGGAGGCCCGGAGGGAAGTGTCTCAGATCCGAAACAGTGAAGAGAGAC	100
QY	301	tacagcgacaagaatgtgaagccagcccatgtgagtcgcagcccgcctcgcgccgcgagag	360
Db	100	TACAGCGCAAAAGTGAAGAGCCAGCGACACTGGAGTGTGCAGCCCGCTCGCGCGCGGGAG	41
QY	361	cgatttcagatttgcgagagagcgccggaagccgataaagaag	400
Db	40	CGGTTTCGAGTTTGGGAGAGCGCCCGGAAGCCAGTAAAGACAG	1

RESULT	6
BG574409	
LOCUS	BG574409
DEFINITION	804 bp mRNA
	602596352P1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4705189 5'
	mRNA sequence.

ACCESSION	BC574409
VERSION	BC574409.1
KEYWORDS	Est.
SOURCE	human.

ORGANISM	Homo sapiens
ENAXYOTA	Eumetazoa: Chordata; Craniata; Vertebrata; Euteleostomi
MAMMALIA	Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 804)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: ccgabs-remail.nih.gov
Tissue Procurement: DCRD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM10571 row: g column: 14
High quality sequence stop: 686.

```

FEATURES
source
location/Qualifiers
1..804
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4705189"
/clone_id="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: breast; Vector: pCMV-SPO6; Site:1: NCI  
Site:2: Salt; Cloned unidirectionally; oligo-dt prim  
Average insert size 1.383 kb. Library enriched for  
full-length clones and constructed by Life Technology  
Note: this is a NIH_MGC library."
BASE COUNT
272 a 165 c 270 g 97 t
ORIGIN

```

Query Match	57.1%;	Score 314.2;	DB 11;	Length 804;
Best Local Similarity	84.5%;	Pred. No. 4.9e-47;		
Matches 388; Conservative	0;	Mismatches 43;	Indels 28;	Gaps

QY	1	caagaacgtcggaaacgaacatctgaagaagaacaggaataacgtccgcgggaattggaaga	60
DB	358	CGAAGACGTGCGGAACGGCATCGAAGAAGAACGATTAAGCTCGCCGGGAATGGGAAGA	411

JOURNAL
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: image.llnl.gov
Plate: L1CM518 row: b column: 07
High quality sequence stop: 667.

FEATURES
Source Location/Qualifiers

1..1042
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3835686"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOT87; Site:1: XhoI; Site:2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 371 a 159 c 390 g 122 t
ORIGIN

Query Match 75.3%; Score 352.6; DB 10; Length 1042;
Best Local Similarity 90.2%; Pred. No. 6,7e-54;

Matches 423; Conservative 0; Mismatches 14; Indels 32; Gaps 3;

Qy 5 aacgttcggaacacatcgagaagaacaggataagctcgcggaatggaagaacaga 64
Db 1 AACGTCGGAACGGCATCGAGAGAACGATTAAGCTCGCGGAATGGGAAGACAGA 60
Qy 65 agagaagaagaatggaag-gagcatccagagagagaaggggaatgatggtgtgc 123
Db 61 AGAGAGGAGATGGCAAGGTCAGCATTCACGAGAGAA----- 99
Qy 124 ctcttcaggaaccgcttggaagcagttagaagaagcggagcggagcgaagatgcg 183
Db 100 -----AGGACCGCTTGAGCAGTTAGAAAGAAAGCGGAGCGGAGCCGAGATGCGG 153
Qy 184 gagcagcagaagcagcggagcagaagaagcggcgcgcgaggagagcggcg 243
Db 154 GAGCAGCAGAGAGACACCGGAGCAGAGAGAGCGCGCGCGCGGAGAGACGGCGC 213
Qy 244 aagagcggagagcccgaggaagtgctcacatcacgaacatgagagagactac 303
Db 214 AAGGAGCGGAGCGCCGAGGAGAGTGTCTCATATCACCAAGATGAGAGACTAC 273
Qy 304 agcgaacaagtgaagccagcactgagtcgcagcccgctcgccgcgcyggaacg 363
Db 274 AGCGCAAAAGTGAAGCCAGCCACTGAGTGCAGCCGCTCGCGCCGCCGCGGAGCGG 333
Qy 364 ttgcaattggagagcggccggaagccagtaa-----aagaagaagaatggaagaagga 419
Db 334 TTCGAGTTGGAGAGCGCCGGAACCAAGTAACGAAGAGAACACTGCGCAAGACAGGA 393
Qy 420 cctgcgtctcgaacttacagacatcacgacgagcgaagcgaagcgaagc 468
Db 394 CCTCTGTGCGACTTACAGGACATCAGCGAGCAGGAGAGAGACACGAC 442

RESULT 4
BC009375 906 bp mRNA HNC 09-JUL-2001
LOCUS BC009375
DEFINITION Homo sapiens, similar to cell division cycle 2-like 1 (PITSLRE proteins), clone IMAGE:4121554, mRNA.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BC009375
BC009375.1 GI:14627288

human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 906)

REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: c9apbs-r@mail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov

Shvchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Barkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-D., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantirpop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
Series: IRAL Plate: 26 Row: d Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 3850303
This clone has the following problem: incomplete processing.

FEATURES
Source Location/Qualifiers

1..906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4121554"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH_MGC_1"
/lab_host="DH10B-R"
/note="Vector: pOT87"
BASE COUNT 317 a 170 c 307 g 112 t
ORIGIN

Query Match 71.9%; Score 336.6; DB 12; Length 906;
Best Local Similarity 91.3%; Pred. No. 4.8e-51;
Matches 379; Conservative 0; Mismatches 9; Indels 27; Gaps 1;

Qy 1 caggaacgttcggaacacatcgagaagaacaggataagctcgcggaatggaagaaga 60
Db 508 CACGAACGTTCGGAACGGCATCGAGAGAACAGATTAAGCTCGCCGGAATGGGAAGA 567
Qy 61 cagaagaagaaggaatggaagaagcattccagaagaagaagaaggggaatgatg 120
Db 568 CAGAAGAAGAGAGATGGAGAGGAGCATTTCCAGAGAGAA----- 609
Qy 121 tgctcttcagggaccgcttggaagcagttagaagaagaagcggagcggcaagatc 180
Db 610 -----AGGACCGCTTGAGCAGTTAGAAAGAGACGGAGCGGAGCGCAAGATG 660
Qy 181 cggagcagcagaagaagcagcggagcagaagaagcggcgcgcgaggagagcgg 240
Db 661 CGGAGCAGCAGAGAGAGACAGCGGAGACAGAAAGAGACCGGAGCGCGGAGAGCGG 720
Qy 241 cgcagaagcgggagcccgaggaagtgtctgcacatcacccaagcgtgagagagac 300

ECORI: cDNA made by oligo-dT priming. Directionally cloned into EcorI/XhoI sites using the following 5', adaptor: GGGACGAG (c). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).

Query Match	84.8%;	Score 396.8;	DB 10;	Length 1063;
Best Local Similarity	93.8%;	Pred. No. 8.5e-62;		
Matches 435; Conservative	0;	Mismatches 2;	Indels 27;	Gaps 1.

QY	5	aagcttggaaacgacatctcgaaagaacaggataaaactcgccgggaaatgtgaaagacaga	64
Db	1	AAGCTTCGAAAACGGCATTCGAAAGAACGAGATTAACCTCGCCGGGAATGGAAACACACA	60
QY	65	agagaaaggaaatctgcaaaggagcatctccagagagaaagaaggggaaatgactgacgtgtgcc	124
Db	61	AGAGAAAGGAGATGTGCGAAGGAGACATTTCCAGGAGAGAA-----	98
QY	125	tcttcaggagaccgctctggagcagttcgaagaagaagccggagccggagacgaagtgcggg	184
Db	99	-----AAGGAGCCGCTGTGACACAGTTTGAAGAAAGGAAGCGGAGCGGAGCCCAAGATGCGGG	153
QY	185	agcagcgaagagagcagccggagcagaagaagacgcgcggccggccggagagacgcgcga	244
Db	154	AGCAGCAGAAAGAGACACACGGGAGACAGAGAGAGCGGCAGCGCGCGGAGAGAGACGGCGCA	213
QY	245	aggagcggagagccgcgcgaggaagtgtctgcacatcacccaagatlgagaagagactaca	304
Db	214	AGGAGCCCGGAGGCCCGCAGGAAAGTGTCTGTCAATCACCGGAAGATGAAGAGAGACTACA	273
QY	305	ggcagcaaaagtgaagaagccagagccactctggaatgcagccgcgcctgcgcgcgcgcggagaggt	364
Db	274	GCGACAAAGTGAAGGCCAGCCACTGTGAGTGTGCAGCCCGCTTCGCGCGCGCGGAGCGGT	333
QY	365	tgcagttggagagacgcgcggagaagccagtaaaagaagaaatlgagaagaagagactgtc	424
Db	334	TGCGATTGGAGAGACGGCCGCGAACCAGTAAATAAAGAAATGGAATGGAAGGAAGGACCTGC	393
QY	425	ttgtcgaacttaacagagcaatcagcgaagaagcgagagagagaccagc	468
Db	394	TGTCGGACTTACAGAGACATTCGCAACGAGCGAGAGGAGACCGAGC	437

RESULT	2
LOCUS	BE742222
DEFINITION	BE742222 723 bp mRNA
ACCESSION	601575562FF1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:383647 5' ,
VERSION	BE742222 mRNA sequence.
KEYWORDS	BE742222 BE742222.1 GI:10156214
SOURCE	EST.
ORGANISM	human.
	Homo sapiens

REFERENCE	1 (bases 1 to 723)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MCC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: LNCM520 row: j column: 08

High quality sequence stop: 702	
FEATURES	Location/qualifiers
source	1. .723

```

/tissue-type="adenocarcinoma cell line"
/lab-host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'-
adaptor: GCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

Query Match	84.6%;	Score 395.8;	DB 10;	Length 723;
Best Local Similarity	93.7%;	Pred. No. 1.3e-61;		
Matches 434;	Conservative 0;	Mismatches 2;	Indels 27;	Gaps 1;

OY		6	aaccggaagacgatcctgaagaagacggataaagctcgccgggaattggaaaagacgaa	65
Dd		1	ACGTGCGAAACGGCATCGAGAAACAAGATTAAGCTTCGCCGGAAATGGAAAACACGAA	60
OY		66	gagaagggaaatigycaaaggagcaatccagtagagaaaaagggaatgatgtgcgttacct	125
Dd		61	GAGAAGGAGATGCGCAAGGGAGCATTCACAGAGAGAA-----	97
OY		126	cttcaggagccgctltgagcagttaagaaggaaagcgggaagcggagacgcgaagtgcggga	185
Dd		98	---AGGCACCCTTTGGAGCGATTGAAGAAAGGAACGGGAGCCGGAGCGCAAGATGCGGGA	153
OY		186	gcacgaagaagagcagcagcgagacgaagaagcgcgagcggcgagcgagagagcgcgcaa	245
Dd		154	GCACAGAAAGSAGCAGCGGGAGCAGAAAGAACGCCGACCGGGCGSGAGSAGCGGCCAA	213
OY		246	ggagcgcggagggcccgcaggggaagtgtctgcacatcacggaacgatatgagagactaacg	305
Dd		214	GGACCGGGAGGCCCCGACGGGAGTGTCTTGCACTCACCGAAGCATGAGAGACTACAG	273
OY		306	cgaacaagtgaanaagcagaagccaatgtcgaagccgccctcgagcgcgcgagggaagcggtt	365
Dd		274	CGACAAAGTGAACCCAKCCACTGTGAATCGAGCCGCCCTCTGGCCGCCGCGGAGCGGTT	333
OY		366	cgaattggagacgcgcggaaagccagttaaagaagaagaaaaatgaaagaagagacgtgct	425
Dd		334	CGAGTTGGGAGACGGCCGGAAAGCCAGTAAAAAGAAAGAAATAATGSAAMAANAAGGAGCTCGCT	393
OY		426	gtccgaacttacagaacatcaagcgacagcgaagagaagacagc	468
Dd		394	GTCGCACTTACAGACATCAACCGCACCGAGAGGAACAACCGAC	436

RESULT	3				
LOCUS	BE742943				
DEFINITION	BE742943	1042 bp	mRNA	EST	15-SEP-2000
ACCESSION	601574225F1		NH_MGC_9	Homo sapiens	cdna clone IMAGE:5835665
VERSION	BE742943.1		GI:10157033		

SOURCE ORGANISM	REFERENCE
human.	1 (bases 1 to 1042)
Homo sapiens	NIH-MGC http://mgc.ncl.nih.gov/ .
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	National Institutes of Health, Mammalian Gene Collection (MGC)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:14:19 : Search time 3881.49 Seconds
(without alignments)
1295.643 Million cell updates/sec

Title: US-09-915-060-5
Perfect score: 468
Sequence: 1 cacgacgctgcgaacgaca.....acagcgagaggaagacacgc 468

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estlom:*
5: em_estlpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hnc:*
10: qd_estcl:*
11: qd_estc2:*
12: qd_hnc:*
13: qd_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_hiv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	396.8	84.6	1063	10	BE740559	BE740559 601595656
2	395.8	84.6	723	10	BE742222	BE742222 601575629
3	352.6	75.3	1042	10	BE742943	BE742943 601574925
4	336.6	71.9	906	12	BC009375	BC009375 Homo sapi
5	332.8	71.1	592	10	AI580044	AI580044 t445908.x
6	314.2	67.1	804	11	BG574409	BG574409 602596352
7	305	65.2	1340	11	BG107702	BG107702 602277878
8	295.2	63.1	1000	12	AK008283	AK008283 Mus muscu
9	281	60.0	662	11	BG822824	BG822824 602727816
10	271.6	58.0	326	11	BG943585	BG943585 ax39f04.x
11	265.8	56.8	826	11	BG034678	BG034678 602300028
12	205.6	43.9	674	10	BE259507	BE259507 601106356

13	203	43.4	372	10	AL034747	AL034747 v8608b53
14	200.4	42.8	812	11	BI149024	BI149024 602911290
15	186.6	39.9	367	10	AA681955	AA681955 v444h10.5
16	172.8	36.9	516	11	BF156008	BF156008 RC0-HT095
17	171	36.5	830	11	BF607177	BF607177 M11-00011
18	162	34.6	758	11	BG969359	BG969359 602836841
19	155.6	33.2	1083	11	BG770237	BG770237 602744821
20	155.2	33.2	493	11	BI018825	BI018825 I13-MT026
21	153	32.7	547	10	BF743770	BF743770 PMO-BT085
22	138	29.5	441	10	AK451883	AK451883 UT-H-B13-
23	127.6	27.3	725	11	BF308197	BF308197 601887445
24	124	26.5	514	10	AI989489	AI989489 ws25d04.x
25	120.8	25.8	462	11	BG486162	BG486162 d6d23e11.
26	120.8	25.8	476	10	BE680738	BE680738 d184d07.Y
27	119.6	25.6	251	10	AV270235	AV270235 AV270235
28	119.4	25.5	732	10	AI006623	AI006623 uc16a10.Y
29	110.2	23.5	160	10	AM193827	AM193827 xm30g12.x
30	106	22.6	673	11	BF983533	BF983533 602306504
31	104.4	22.3	469	11	BF359805	BF359805 RC6-MT006
32	103.8	22.2	321	10	AA309711	AA309711 EST180769
33	102.4	21.9	233	10	BB290670	BB290670 BB290670
34	100.4	21.5	253	10	AV327502	AV327502 AV327502
35	96.4	20.6	372	11	BF918032	BF918032 I13-UT011
36	95	20.3	292	10	AV368285	AV368285 AV368285
37	86.8	18.5	479	11	BF934664	BF934664 I12-NT020
38	85.2	18.2	253	10	AV331446	AV331446 AV331446
39	85	18.2	182	10	AA13247	AA13247 v652d12.T
40	83.8	17.9	702	10	BE308367	BE308367 601090048
41	80.4	17.2	202	10	AV330095	AV330095 AV330095
42	76.4	16.3	987	13	CNS00418	AL066537 Dicosoph11
43	74.4	15.9	997	10	CNS005TE	AL066537 Dicosoph11
44	73	15.6	740	10	AL043340	AL043340 DKF2P340
45	71.8	15.3	912	13	CNS02LBC	AL202705 Tetradon

ALIGNMENTS

RESULT 1
LOCUS BE740559 1063 bp mRNA
DEFINITION 601595656F1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3949693 5',
MRNA sequence.
ACCESSION BE740559
VERSION BE740559.1 GI:10154551
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: cgrabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DMF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: L1CM814 row: p column: 14
High quality sequence stop: 741.
Location/Qualifiers
1..1063
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949693"
/clone_l1b="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:

FEATURES

source

LOCATION: 1507.1644
FEATURE:
NAME/KEY: inton
LOCATION: 1645.2511
FEATURE:
NAME/KEY: CDS
LOCATION: 2512.8070
US-08-800-644-93

Query Match 12.2%; Score 57.2; DB 2; Length 9551;
Best Local Similarity 51.3%; Pred. No. 0.00017;
Matches 159; Conservative 0; Mismatches 148; Indels 3; Gaps 1;

QY 54 ggaagacagaagaagaagaatgacagagagatccagagagaagaagaatga 113
DB 3192 ggaagacagaagaagaagaatgacagagagatccagagagaagaaga 3251
QY 114 tggcgtgtgtctcttcacagcgccttgagcagttagaagaagcagcagcag 173
DB 3252 acagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3308
QY 174 caagatgcagagacagaagaagaagaagaagaagaagaagaagaagaaga 233
DB 3309 ggaagacagaagaagaagaagaagaagaagaagaagaagaagaagaaga 3368
QY 234 ggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 293
DB 3369 ggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3428
QY 294 aagaagacacagcagcagcagcagcagcagcagcagcagcagcagcagcag 353
DB 3429 ggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3488
QY 354 gcgagcagcag 363
DB 3489 ggcgcagcagcag 3498

RESULT 14
US-08-781-891-208/c
Sequence 208, Application US/08781891
Patent No. 6090630
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-Hui
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620Lemburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 16442 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-208

Query Match 11.9%; Score 55.6; DB 3; Length 16442;
Best Local Similarity 65.1%; Pred. No. 0.00042;
Matches 82; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 141 ggaacagttagaagaagaagaagaagaagaagaagaagaagaagaagaaga 200
DB 16376 ggaacagttagaagaagaagaagaagaagaagaagaagaagaagaaga 16317
QY 201 gcgagacagaagaagaagaagaagaagaagaagaagaagaagaagaaga 260
DB 16316 ggaacagttagaagaagaagaagaagaagaagaagaagaagaagaaga 16257
QY 261 caagga 266
DB 16256 ggaaga 16251

RESULT 15
US-09-253-691-3
Sequence 3, Application US/09253691
Patent No. 6124100
GENERAL INFORMATION:
APPLICANT: Dong Kyu JIN
TITLE OF INVENTION: Diagnostic Method and Kit for Neuropsychiatric Diseases
FILE REFERENCE: 1942/36
CURRENT FILING DATE: 1999-02-22
EARLIER APPLICATION NUMBER: KR 98-6,278
EARLIER FILING DATE: 1996-02-26
NUMBER OF SEQ ID NOS: 3
SOFTWARE: WordPerfect 6.1/Windows
SEQ ID NO 3
LENGTH: 397
TYPE: DNA
ORGANISM: human
US-09-253-691-3

Query Match 11.7%; Score 54.8; DB 3; Length 397;
Best Local Similarity 62.3%; Pred. No. 0.00037;
Matches 86; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 141 ggaacagttagaagaagaagaagaagaagaagaagaagaagaagaagaaga 200
DB 235 ggaacagttagaagaagaagaagaagaagaagaagaagaagaagaagaaga 294
QY 201 gcgagacagaagaagaagaagaagaagaagaagaagaagaagaagaaga 260
DB 295 ggaacagttagaagaagaagaagaagaagaagaagaagaagaagaagaaga 354
QY 261 caaggaagtgtctgcaca 278
DB 355 gcgagacattatcagcaca 372

Search completed: December 4, 2001, 18:57:53
Job time: 9219 sec

Db 2911 GGCAGAGCAGGAGGGCGAGAGCAGAGGTGGAGCCGGGCTCAGAGAGCGAGTGAGAG 2970
QY 320 ccag 323
Db 2971 CCGG 2974

RESULT 10
US-08-194-087-15
; Sequence 15, Application US/08194087
; Patent No. 5879910
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,087
; FILING DATE: 18-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-194-087-15

Query Match 12.3%; Score 57.6; DB 2; Length 10596;
Best Local Similarity 49.3%; Pred. No. 0.00014;
Matches 150; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY 20 atcgaagaagaacagataaagctcgcggaatggaagaagaagaagaagaatgg 79
Db 2671 ACCAGAGAGGCGCAGAGAGAGGCGCAGAGAGGCGCAGAGAGGCGCAGAGAGG 2730
QY 80 caaggagagatcccaaggagagaagaagggaatgatgagcgtgtgtcctcttcagggaaccgt 139
Db 2731 AGGGGAGAGAGGCGCAGAGAGAGGAGGAGGAGGCGCAGAGAGAGAGAGAGAGGCGCAGG 2790
QY 140 tggagcagttagaagaagaagcggagcggagcggagcgaagatgcggagagcagaagaagagc 199
Db 2791 AGGGGAGAGAGAGAGAGAGGAGGCGCAGAGAGGCGCAGAGAGAGGAGGCGCAGG 2850
QY 200 agcggagagcagaagaagcggagcggagcggagcggagcggagcggagcggagcggagcggc 259
Db 2851 ACCAGAGAGGCGCAGAGAGGCGCAGAGAGGAGGAGGCGCAGAGAGGCGCAGAGAGAGG 2910
QY 260 gcaagggaagtgctctgcaatcaaccgaacgatgtgagagagagactaacgcgcgcaaaagtgaag 319

Db 2911 GGCAGAGCAGGAGGGCGCAGAGAGCGGTGGAGCCGGGCTCAGAGAGCGAGTGAGAG 2970
QY 320 ccag 323
Db 2971 CCGG 2974

RESULT 11
PCT-US93-04648-15
; Sequence 15, Application PC/TUS9304648
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc., Godowski, Paul J., Lokker, Nathalie A., Mark, Melani
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/04648
; FILING DATE: 19930517
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 755,779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US93-04648-15

Query Match 12.3%; Score 57.6; DB 5; Length 10596;
Best Local Similarity 49.3%; Pred. No. 0.00014;
Matches 150; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
QY 20 atcgaagaagaacagataaagctcgcggaatggaagaagaagaagaagaatgg 79
Db 2671 AGCGAGAGGCGCAGAGAGAGGCGCAGAGAGGCGCAGAGAGGCGCAGAGAGG 2730
QY 80 caaggagacattcccaaggagagaagaagggaatgatgagcgtgtgtcctcttcagggaaccgt 139
Db 2731 AGGGGAGAGAGGCGCAGAGAGAGGAGGCGCAGAGAGGCGCAGAGAGAGAGAGAGGCGCAGG 2790
QY 140 tggagcagttagaagaagaagcggagcggagcggagcgaagatgcggagagcagaagaagagc 199
Db 2791 AGGGGAGAGAGAGAGAGAGGAGGCGCAGAGAGGCGCAGAGAGGCGCAGAGAGGCGCAGG 2850
QY 200 agcggagagcagaagaagcggagcggagcggagcggagcggagcggagcggagcggagcggc 259
Db 2851 AGCGAGAGGCGCAGAGAGGCGCAGAGAGGAGGCGCAGAGAGGCGCAGAGAGAGGAGGAGG 2910

Qy	320	ccag	323
Db	1437	CCGG	1440

RESULT 6

US-07-884-811-15
Sequence 15, Application US/07884811
Patent No. 5316921
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCTE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: palin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/884,811
FILING DATE: 19920518
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dieger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 755.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/425-3316
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 10596 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: Single
TOPOLOGY: linear
US-07-884-811-15

Query Match	12.3%	Score 57.6	DB 1	Length 10596
Best Local Similarly	49.3%	Pred. No. 0.00014		
Matches 150; Conservative	0;	Mismatches 154;	Indels 0;	Gaps 0;

[illegible]

QY	320	ccag	323
Db	2971	CCGG	2974

RESULT

```

US-07-885-971-15
/ Sequence 15, Application US/07885971
/ Patent No. 5328837
/ GENERAL INFORMATION:
/ APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
/ TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Genentech, Inc.
/ STREET: 460 Point San Bruno Blvd
/ CITY: South San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94080
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: patin (Genentech)
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/885,971
/ FILING DATE: 19920518
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Dreger, Ginger R.
/ REGISTRATION NUMBER: 33,055
/ REFERENCE/DOCKET NUMBER: 779
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415/225-3216
/ TELEFAX: 415/952-9881
/
/ TELEX: 910/371-7168
/ INFORMATION FOR SEQ ID NO: 15:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1056 bases
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-07-885-971-15

```

Query Match	12.38;	Score 57.6;	DB 1;	Length 10596;
Best Local Similarity	49.38;	Pred. No. 0.00014;		
Matches 150; Conservative	0;	Mismatches 154;	Indels 0;	Gaps 0;

Accession	Sequence	Length
QY	20 atctgaagaagaacaaagataaagcttcgcccggaataatgggaaagaaacaaagaaagaaaggaatg	79
Db	2671 AGCAGAGAGGGGCAGGAGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGG	2730
QY	80 caaaggagcattccacagagagaaagggggaatatgatgcytctgctcttcacagagacgct	139
Db	2731 AGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGG	2790
QY	140 ttggagcagttatgaagaagaagcgggaagcgggaagcgaagatctcggaagcagcagaagaagac	199
Db	2791 AGGGGCAGGAGCAGGAGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGG	2850
QY	200 agcgggaagcagaagaagcgcgagcggcgggcggaaggaagcgcgcaagggagcgggaagcc	259
Db	2851 AGCAGAGGGGCAGGAGGGGCAGGAGCAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGG	2910
QY	260 gcaaggaaagtgtctgcatacatccgaacgataagagagagtaacagcgaacaagtgaag	319
Db	2911 GGCAGAGCAGGAGGGGCAGGAGCAGGAGGTGAGGACCGGGGTCGAGAGGAGCAGGAGGAGG	2970

Query Match	12.3%	Score 57.6	DB 3	Length 2580
Best Local Similarity	49.3%	Pred. No. 0.00012		
Matches 150	Conservative	0	Mismatches 154	Indels 0
QY 20	atcgaagaacagatataagctcgcgcggaatctggaagaacagaagaagaaatg	79		
Db 1091	AGCAGAGAGGGCGACGAGCAGCGAGGGCGAGCAGCAGAGAGGGCGACGAGCAGG	1150		
QY 80	caagagagcatctccaggaagaagaaggaatgatgctgtgtcctcttcaggaacgct	139		
Db 1151	AGGGCGAGAGGGCGAGCAGGAGGGCGAGAGGGCGCAGGACGACGAGAGGGCGAGG	1210		
QY 140	tggagcagtttaagaagagagcgcgagcgcggaagcgaagatgcgggaagcagcagaagagc	199		
Db 1211	AGGGCGAGAGCAGCAGGAGGAGGGCGAGGAGGGCGAGCAGCAGCAGAGAGGGCGAGG	1270		
QY 200	agcggagacagaagagcgcgagcgcgcggaagagcgcgcaagagagcgaggaagccc	259		
Db 1271	AGCAGAGAGGGCGACGAGAGGGCGACGAGAGGGCGACGAGAGGGCGACGAGAGGAGG	1330		
QY 260	gcaaggaagtgtctcacatccagcaacgatatgagagagatctacagcgaacagtgaag	319		
Db 1331	GGCAGAGAGCAGGAGGGCGAGGAGGTGAGAGGCCGGGTGAGGAGGCGAGTGTGAGG	1390		
QY 320	ccag 323			
Db 1391	CCGG 1394			

```

RESULT      4
US-09-130-114-1/c
: Sequence 1, Application US/09130114
: Patent No. 5976807
: GENERAL INFORMATION:
: APPLICANT: Horlick, Robert A.
: APPLICANT: Dama, Bassam B.
: APPLICANT: Robbins, Alan K.
: TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
: TITLE OF INVENTION: From Multiple Transfected Episomes
: FILE REFERENCE: 0867/1D903US1
: CURRENT APPLICATION NUMBER: US/09/130,114
: CURRENT FILING DATE: 1998-08-06
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 3452
: TYPE: DNA
: ORGANISM: HBVNA
US-09-130-114-1

```

	Query Match	12.3%	Score 57.6	DB 2	Length 5452
	Best Local Similarity	49.3%	Pred. NO.0.0013		
	Matches 150	Conservative	0	Mismatches 154	Indels 0
					Gaps 0
QY	20	atcgaagaacacagatataaagctcgcgcggaattggaaagacagaagaagaagaaatg	79		
Db	1714	AGCAGAGGGGCGAGCAGCAGAGGGGCGAGCAGCAGAGGGGCGAGGCGCAGAGCAGG	1655		
QY	80	caagggaacatccacgaagaagaagggaatgatgtgcgtgtcccttcaaggacgct	139		
Db	1654	AGGGCGAGGAGGGGCGAGCAGCAGGAGGGGCGAGGAGGGGCGAGCAGCAGGAGGAGG	1595		
QY	140	tggagcaatttgaagaagacgagagcgagagcggaatgcggaagcagcagaagagac	199		
Db	1594	AGGGCGAGCAGCAGGAGGAGGGGCGAGGAGGGGCGAGCAGCAGCAGGAGGGGCGAGG	1535		
QY	200	agcgggaacagaagagcgagcgcgcgcggaagaagcgcgcagaagacgagagagccc	259		
Db	1534	AGCAGAGGGGCGAGGAGGGGCGAGGAGCAGGAGGGGCGAGGAGGCGAGGAGGAGG	1475		
QY	260	gcagggaagtgtctgcacatcaacggaacgatagagagactacagcgcacaagtgaag	319		

Dd	1474	GGCAGGACGAGAGGGCAGACAGGAGGTGAGGCCGGGGTCGAGGAGGCAGTGGAGG	1415
QY	320	ccag 323	
Dd	1414	CCGG 1411	

```

RESULT      5
US-08-910-647-1
: Sequence 1, Application US/08910647
: Patent No. 6251433
: GENERAL INFORMATION:
: APPLICANT: Zuckermann et al.
: TITLE OF INVENTION: Compositions and Methods for
: TITLE OF INVENTION: Polynucleotide Delivery
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation
: STREET: 4560 Horton Street
: CITY: Emeryville
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 94608-2916
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/910,647
: FILING DATE:
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Fujita, Sharon M.
: REGISTRATION NUMBER: 38,459
: REFERENCE/DOCKET NUMBER: 1218.002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 923-2706
: TELEFAX: (510) 655-3542
: INFORMATION FOR SEQ. ID NO.: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 9600 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-08-910-647-1

```

Query Match	12.3%	Score 57.6	DB 4	Length 9600
Best Local Similarity	49.3%	Pred. No. 0	0014	
Matches 150	Conservative 0	Mismatches 134	Indels 0	Gaps 0
QY	20	atcagaagaacacagataaagctcgccggygaatgygaagaacagaagaaggaatg	79	
Db	1137	AGCAGAGAGGGGCGAGAGAGAGGGGGCGAGAGACAGAGAGGGGCGAGAGGCGCAGGAGCAGG	1196	
QY	80	caaggaagcattcccaaggaagaagaaggggaatgtatgagctgtgctcttcaggaacgct	139	
Db	1197	AGGGGCGAGGAGGGCGAGGACGAGAGGGGGCGAGAGGGGCGACGAGCAGGACGAGAGGGGCGAGG	1256	
QY	140	tygsagcagttagaagaaggaagcgggagcgggagcgaagatgcyggagcagcagaaggaagc	199	
Db	1257	AGGGGCGAGAGCAGAGAGAGGGGGCGAGGAGGCGGCGAGACAGAGAGGGCGCGAGAGGGCGCAGG	1316	
QY	200	aagcggaacaaagaagagcgagcggcggggggagacggcgcaagagacgggaagggcc	259	
Db	1317	AGCAGAGAGGGGCAAGAGAGGGCGAGGACAGGAGAGGGGCGAAGAGAGGGCGAGGAGGAGG	1376	
QY	260	gcaggaagatgtcttcacatcaccgnaagatgatagagagactacagcgacaagaagtgaag	319	
Db	1377	GGCAGAGACAGAGAGGGGCGAGGACAGGAGGTGTGAGAGCGGGGTGTGAGGACGAGGAGG	1436	

```
Query Match      21.2%; Score 99.4; DB 1; Length 7218;
Best Local Similarity 5.4%; Pred. No. 7.3e-14;
Matches 22; Conservative 259; Mismatches 130; Indels 0; Gaps 0;

Qy 2 acgaacgtcggaacgacatcggaagaacagataagctgcgcggatggaagac 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1469 ATGCACTACTTTAAAGAGATAGAGATTTCGTACRRRRRRRRRRRRRRRR 1410
Qy 62 agaaagaaaggaatgcaagagacatccagagagaagaaaggaatgagcgtg 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1409 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1350
Qy 122 gctcttcggggcgccgtgagcagltgaaagaaagcgggagcgagcgcaagatgc 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1349 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1290
Qy 182 ggaagacagaaagagcgagcgagcaagaagagcgagcgggcggaagcgagc 241
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1289 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1230
Qy 242 gcaagagcgggagcgccgcaaggagtgctgcacatccagacatgagagagact 301
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1229 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1170
Qy 302 acagcgacaaagtgaagcgccagcactgagtcgagcgccgctcgccgcgagagc 361
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1169 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1110
Qy 362 ggttcgagtgaggagcgccgagcgcaataaagaagaatgagag 412
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1109 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1059

RESULT 2
US-08-458-568A-11/c
; Sequence 11, Application US/08458568A
; Patent No. 5821339
; GENERAL INFORMATION:
; APPLICANT: Schaffer, Priscilla A.
; TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5621339rits
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,568A
; FILING DATE: 02-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/065,146
; FILING DATE: 05-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn R.
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: DFCI-0029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12001 base pairs
```

```
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Herpes simplex virus
STRAIN: Herpes Simplex Virus Type 1
US-08-458-568A-11

Query Match      12.4%; Score 58; DB 1; Length 12001;
Best Local Similarity 57.9%; Pred. No. 0.00012;
Matches 103; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 79 gcaagggagcattccagagagaaaggggaatgagcgtgtgctcttcagggacgc 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1541 GCGGGCGCGCGCGCGCGCAAAAAGCGCGCGCGCGCGCGCGCGCGCGCG 1482
Qy 139 ttgagcagttagaagaaagcgagcgagcgcaagatgagcgagcgagaaagag 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1481 CGGGCGTGGGGCGCGGCGCCCGGAGCGGGGAGAGACCGGGGAGAGCGGGGAG 1422
Qy 199 cagcgagagcagaagagcgagcgagcgagcgagcgagcgagcgagcgagag 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1421 GAGCGGGGAGAGAGCGGGGAGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAG 1364

RESULT 3
US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-050-863-2
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:57:33 ; Search time 152.61 Seconds
(without alignments)
694,527 Million cell updates/sec

Title: US-09-915-060-5

Perfect score: 468
Sequence: 1 caccagacgtcggaacgaca.....acagcgagagagagaccagc 468

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/1na/5A_COMB.seq:*
2: /cgn2_6/prodata/2/1na/5B_COMB.seq:*
3: /cgn2_6/prodata/2/1na/6A_COMB.seq:*
4: /cgn2_6/prodata/2/1na/6B_COMB.seq:*
5: /cgn2_6/prodata/2/1na/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/1na/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	99.4	21.2	7218	1	US-08-232-463-14
C 2	58	12.4	12001	1	US-08-458-568A-11
C 3	57.6	12.3	2580	3	US-09-050-863-2
C 4	57.6	12.3	5452	2	US-09-130-114-1
C 5	57.6	12.3	9600	4	US-08-910-647-1
6	57.6	12.3	10596	1	US-07-884-811-15
7	57.6	12.3	10596	1	US-07-885-971-15
8	57.6	12.3	10596	1	US-08-087-783A-15
9	57.6	12.3	10596	1	US-08-194-088B-15
10	57.6	12.3	10596	2	US-08-194-087-15
11	57.6	12.3	10596	5	PCT-US93-04648-15
12	57.2	12.2	9551	1	US-08-056-200-93
13	57.2	12.2	9551	2	US-08-800-644-93
14	55.6	11.9	16442	3	US-08-781-891-208
15	54.8	11.7	397	3	US-09-253-691-3
C 16	54	11.5	51259	3	US-08-781-891-209
17	53.8	11.5	2301	5	US-08-306-691B-23
18	53.8	11.5	2301	5	PCT-US93-06251-78
19	53	11.3	289	4	US-09-007-005-17
20	53	11.3	289	4	US-09-244-796-17
21	51.4	11.0	168	1	US-08-469-802B-4
22	51.4	11.0	168	2	US-08-267-803B-4
23	51.4	11.0	171	1	US-08-469-802B-5
24	51.4	11.0	171	2	US-08-267-803B-5
25	51.4	11.0	195	1	US-08-469-802B-2
26	51.4	11.0	195	2	US-08-267-803B-2
27	51.4	11.0	234	1	US-08-469-802B-3

28	51.4	11.0	234	2	US-08-267-803B-3
29	51	10.9	3376	1	US-08-320-559-29
30	51	10.9	3376	3	US-08-545-660D-29
31	51	10.9	3376	5	PCT-US94-04496-29
32	50	10.7	203	4	US-09-043-303-7
33	49.4	10.6	477	4	US-09-135-994-1
34	49.2	10.5	3489	2	US-08-728-323A-1
C 35	49.2	10.5	32207	2	US-08-770-379-20
C 36	49.2	10.5	32207	4	US-08-757-669A-20
37	48.4	10.3	154	1	US-08-469-802B-6
38	48.4	10.3	154	2	US-08-267-803B-6
C 39	48.2	10.3	1931	2	US-09-130-114-2
C 40	47.8	10.2	165	4	US-09-043-303-17
41	45.4	9.7	543	6	5273901-6
C 42	44.8	9.6	53526	3	US-08-658-136-2
C 43	44.8	9.6	53577	3	US-08-658-136-1
44	44.6	9.5	250	1	US-08-480-552-16
45	44.6	9.5	250	3	US-08-929-208-16

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)883-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-Fls
US-08-232-463-14

Sequence 3, Appl1
Sequence 29, Appl1
Sequence 29, Appl1
Sequence 29, Appl1
Sequence 1, Appl1
Sequence 1, Appl1
Sequence 17, Appl1
Sequence 20, Appl1
Sequence 20, Appl1
Sequence 6, Appl1
Sequence 6, Appl1
Sequence 2, Appl1
Sequence 17, Appl1
Patent No. 5273901
Sequence 2, Appl1
Sequence 1, Appl1
Sequence 16, Appl1


```
Search completed: December 4, 2001, 19:03:42
Job time: 6983 sec
```

XX	Epstein-Barr virus.
XS	
XX	W09822577-A1.
PN	
XX	28-MAY-1998.
PD	
XX	17-NOV-1997; 97MO-IB01508.
PE	
XX	25-JUN-1997; 97US-0048945.
PR	15-JUN-1996; 96US-0030986.
PR	
XX	(MASU/) MASUCI M. G.
PA	
XX	Masucci MG;
P1	
DR	WPI; 1998-312463/27.
XX	
PT	New fusion proteins resistant to proteolytic degradation -
PT	comprising a core protein with a stabilising polypeptide comprising
PT	a peptide sequence containing glycine repeats
PS	
XX	Disclosure; Fig 4B; 120pp; English.
XX	
XX	This is a nucleotide sequence of the stabilising sequence-encoding
CC	insert. The invention provides a method for increasing the resistance
CC	of a core protein or to proteolytic degradation that comprises linking or
CC	inserting onto or into the core protein a stabilising polypeptide of
CC	formula [(Gly)xX(Gly)yX(Gly)z]n where Glya, Glyb, Glyc are 1-6
CC	sequential gly residues and x, y, z are Ala, Ser, Val, Ile, Leu, Met,
CC	Phe, Pro or Thr and n can be anything between 1-66. X, Y and Z need not
CC	be identical from n repeat to n repeat. Alternatively a nucleic acid
CC	encoding the stabilising polypeptide can be linked onto or inserted into
CC	a nucleic acid encoding a core protein. The fusion proteins of the
CC	invention are more resistant to degradation by proteases and, thus, have
CC	a longer half-life than the unused core protein. The products can be
CC	used for treating autoimmune diseases, cancer and inflammation. In
CC	particular, the core protein may be an Ikappa regulator protein for the
CC	treatment of inflammatory bowel disease, or a nitroductase protein
CC	which can activate nitro drugs in enzyme/prodrug therapy to treat cancer
CC	or other pathological conditions. The fusion proteins can also be used in
CC	diagnostic methods such as in vivo imaging.
QX	
QX	Sequence 799 BP; 201 A; 106 C; 479 G; 13 T; 0 other;

DR WPI: 1995-022825/03.
XX Herpes Simplex Virus (HSV) specific junction spanning transcript
PT - for inhibiting HSV L/ST synthesis, in the treatment of HSV
PT Infection.
XX
XX
PS Disclosure; Page 38-44; 64pp; English.
XX
XX An HSV-specific junction-spanning transcript (L/ST) maps at the 5'
CC end to the b repeat sequences of HSV DNA at approx. 3 and 125 kb,
CC and extends into the c repeat sequences of HSV DNA. Compounds
CC that inhibit HSV L/ST synthesis may be used as anti-HSV
CC virucides. The HSV-1 DNA sequence in the region of the L/STs
CC is given in AA076213.
XX
XX
SQ Sequence 12001 BP; 1568 A; 4557 C; 4315 G; 1561 T; 0 other;

Query Match 12.4%; Score 58; DB 16; Length 12001;
Best Local Similarity 57.9%; Pred. No. 0.00021;
Matches 103; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 79 gcaaggagacattccagagagagaggaatgagtgtgtgtctcttcagggaccgc 138
DB 1541 GCGGGCGCGCGCGCGCAAAAAGCGCGCGCGCTCCGCGCGCGCGCGCGCG 1482
QY 139 ttgagagcagttagaagaagcagcagcagcagcagcagcagcagcagcagc 198
DB 1481 CGGGCGTGGGGGGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1422
QY 199 cagcggagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 256
DB 1421 GAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGCGGGG 1364

RESULT 14
AAD07132
ID AAD07132 standard; cDNA; 2803 BP.
XX
XX
AC AAD07132;
XX
DT 06-AUG-2001 (first entry)
XX
DE Canine retinitis pigmentosa GTPase regulator (RPGR) mutant #2 cDNA.
XX
KW Dog: X-linked progressive retinal atrophy 2; XLPRA2; genetic marker;
KW retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
KW Miniature Schnauzer; mutant; muteln; ss.
XX
OS Canis familiaris.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
CDS 1..1149
FT /*tag= a
FT /*product= "Canine retinitis pigmentosa GTPase
FT regulator mutant"
FT /note= "CDS does not include start codon"
FT /partial
FT replace (931..932, AGAG)
FT /*tag= b
FT /note= "This deletion results in the change of
FT amino acids and ends in a premature stop codon"
XX
XX
PN M0200138578-A1.
XX
XX
PD 31-MAY-2001.
XX
XX
PF 21-NOV-2000; 2000MO-US31940.
XX
PR 24-NOV-1999; 99US-0167365.
XX
XX
PA (CORR) CORNELL RES FOUND INC.

XX
PI Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
XX WPI: 2001-367707/38.
DR P-PSDB; AAE02399.
XX
XX
PT Identifying dogs with or carrying X-linked progressive retinal atrophy
PT by detecting retinitis pigmentosa GTPase regulator gene mutation,
PT useful when breeding Husky, Samoyed and Miniature Schnauzer -
XX
XX
PS Claim 75; Page 34-35; 88pp; English.
XX
XX
XX The invention relates to a method for identifying dogs which are
CC genetically normal, are carriers of, or are affected with X-linked
CC progressive retinal atrophy (XLPRA), by testing a biological sample with
CC genetic markers that co-segregate with a XLPRA gene locus. The invention
CC also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
CC mutants and their corresponding nucleic acid molecules. The mutated RPGR
CC genes are responsible for the XLPRA in dogs. Methods are used to select
CC dogs for breeding so that dogs carrying the mutated locus are eliminated
CC from the breeding stock. The method particularly applies to Siberian
CC Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
CC the disease is X-linked. XLPRA1 type is identified in Siberian Huskies,
CC and Samoyeds, while XLPRA2 type is identified in Miniature Schnauzers.
CC The present cDNA sequence is the mutant open reading frame (ORF) 15
CC encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant
CC 'GA' nucleotide bases from position 932 to 933 of the wild-type
CC canine RPGR cDNA.
XX
XX
SQ Sequence 2803 BP; 1028 A; 287 C; 939 G; 549 T; 0 other;

Query Match 12.4%; Score 57.8; DB 22; Length 2803;
Best Local Similarity 51.8%; Pred. No. 0.00017;
Matches 131; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 47 ggaagaggagacacagcagagaggaatgtgcagagagacattccagagagagag 106
DB 913 gaggaaagagaaatgaaatgaaagagagaaagaaagagagagagagagagag 972
QY 107 ggaatgagtggt 166
DB 973 ggaagagagagtggt 1032
QY 167 gggagagcagagtggt 226
DB 1033 agggagagagagtggt 1092
QY 227 gggcggagagagcggcgcaagagcgggagccgagggagagtgctcacatccgaa 286
DB 1093 ggggggggagagggagagagagggagggagagagagagagagagagagagag 1152
QY 287 cgatgagagagga 299
DB 1153 ggaagagagagta 1165

RESULT 15
AAV55831
ID AAV55831 standard; DNA; 799 BP.
XX
XX
AC AAV55831;
XX
XX
DT 18-NOV-1998 (first entry)
XX
XX
DE Nucleotide sequence of the stabilising sequence-encoding insert.
XX
XX
KW Fusion protein: stabilising polypeptide; proteolytic degradation;
KW resistance: half-life; autoimmune disease; inflammation; nitro drug;
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
KW cancer; pathological condition; ss.

DB	441	gcgagcgag	450
<hr/>			
RESULT 10			
ID	AAAI2156	standard; DNA; 2875 BP.	
XX	AAAI2156;		
AC	AAA12156;		
XX	10-AUG-2000	(first entry)	
DE	Human X5L DNA fragment.		
DE	X5L; human; spermatogenesis protein; treatment; diagnosis; detection; ds.		
KW	Homo sapiens.		
OS	DEJ9856882-CL1.		
PN	04-MAY-2000.		
PD	10-DEC-1998;	98DE-1056882.	
PE	10-DEC-1998;	98DE-1056882.	
XX	10-DEC-1998;	98DE-1056882.	
PR	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.		
XX	Sedlacek Z, Poustka A;		
PI	WPf: 2000-294180/26.		
XX	P-PDSB; AAAI2156.		
DR	New spermatogenesis protein useful for investigation and modulation of		
PT	spermatogenesis, particularly for diagnosis and treatment of		
PT	spermatogenesis disorders -		
XX	Claim 4; Fig 2; 14pp: German.		
PS	This invention describes a novel spermatogenesis protein (I) and the DNA		
CC	molecule (II) encoding (I). (I) and (II) are used to investigate and		
CC	modulate spermatogenesis, especially to activate or inhibit it and for		
CC	diagnosis and treatment of spermatogenesis disorders. (I) is also used		
CC	to raise specific antibodies, useful for detection of (I) and as		
CC	therapeutic inhibitors of (I), and to detect specific autoantibodies. (II)		
CC	can be used as a source of primers to determine the organization and		
CC	expression of the (I)-encoding gene, also for in vivo expression of (I).		
CC	This sequence encodes the human spermatogenesis protein X5L which is		
CC	described in the method of the invention.		
XX			
SQ	Sequence 2875 BP; 616 A; 797 C; 894 G; 568 T; 0 other;		
<hr/>			
Query Match 12.7%; Score 59.6; DB 21; Length 2875;			
Best Local Similarity 66.2%; Pred. No. 6.6e-05;			
Matches 86; Conservative 0; Mismatches 44; Indels 0; Gaps 0;			
OY	117	Cygtgtcctttaaaggaccgtttgaagcaatagaagaacggcggaacgcgcga	176
DG	1359	cctgaacataatattatatttttttttttttttttttttttttttttttttt	1418
OY	177	gatcgcgagcacacaagaagacagcgagacagagagcgagcgcgcgagagga	236
DG	1419	caagcgccagcactcttgaaagacagcgcgtctgcagcagagcgcgagacagga	1478
OY	237	gcgcgcgcaag	246
DG	1479	gcgcgcgcaag	1488
<hr/>			
RESULT 11			
ID	AAD07131	standard; cDNA; 2800 BP.	
XX	AAD07131 standard; cDNA; 2800 BP.		

AC AAD07131;
XX
DT 06-AUG-2001 (first entry)
XX
DE Canine retinitis pigmentosa GTPase regulator (RPGR) mutant #1 cDNA.
XX
KW Dog; X-linked progressive retinal atrophy 1; XLPRAL; genetic marker;
XX retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
KM Miniature Schnauzer; mutant; mutein; ss.
OS
OS Canis familiaris.
XX Synthetic.
FH
FH
FH Key Location/Qualifiers
CDS 1..882
FT /*tag= a
FT /product= "Canine retinitis pigmentosa GTPase
FT regulator mutant"
FT /note= "CDS does not include start codon"
FT /partial
FT mutation replace (877..878, AGAGAAAT)
FT /*tag= b
FT /note= "This deletion replaces Arg with Ile and results
in a premature stop codon"
PN WO200138578-A1.
XX
PD 31-MAY-2001.
XX
PE 21-NOV-2000; 2000WO-US31940.
XX
PR 24-NOV-1999; 99US-0167365.
XX
PA (CORR) CORNELL RES FOUND INC.
XX
XX Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
PL WPI; 2001-367707/38.
DR P-PSDB; AAEO2398.
XX
PT Identifying dogs with or carrying X-linked progressive retinal atrophy
PT by detecting retinitis pigmentosa GTPase regulator gene mutation,
PT useful when breeding Husky, Samoyed and Miniature Schnauzer -
XX
XX
PS Claim 73; Page 32-33; 88pp; English.
XX
CC The invention relates to a method for identifying dogs which are
CC genetically normal, are carriers of, or are affected with X-linked
CC progressive retinal atrophy (XLPRAL), by testing a biological sample with
CC genetic markers that co-segregate with a XLPRAL gene locus. The invention
CC also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
CC mutants and their corresponding nucleic acid molecules. The mutated RPGR
CC genes are responsible for the XLPRAL in dogs. Methods are used to select
CC dogs for breeding so that dogs carrying the mutated locus are eliminated
CC from the breeding stock. The method particularly applies to Siberian
CC Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
CC the disease is X-linked. XLPRAL type is identified in Siberian Huskies,
CC and Samoyeds, while XLPRAL2 type is identified in Miniature Schnauzers.
CC The present cDNA sequence is the mutant open reading frame (ORF) 15
CC encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant
CC found in XLPRAL-affected dogs. This mutant is obtained by deleting
CC 'GAGA' nucleotide bases from position 878 to 882 of the wild-type
CC canine RPGR cDNA.
CC
SQ Sequence 2800 BP; 1026 A; 287 C; 938 G; 549 T; 0 other;

Query Match 12.7%; Score 59.4; DB 22; Length 2800;
Best Local Similarity 52.2%; Pred. No. 7.3e+05;
Matches 132; Conservative 0; Mismatches 121; Indels 0; Caps 0;

47 gggagatggaaagacgaagaagggaagtgcgaaggagcattccagagagaaggg 106

CC used for treating autoimmune diseases, cancer and inflammation. In
CC particular, the core protein may be an IkappaB regulator protein for the
CC treatment of inflammatory bowel disease, or a nitroreductase protein
CC which can activate nitro drugs in enzyme/prodrug therapy to treat cancer
CC or other pathological conditions. The fusion proteins can also be used in
CC diagnostic methods such as in vivo imaging.

XX Sequence 795 BP; 200 A; 104 C; 478 G; 13 T; 0 other;

Query Match 13.2%; Score 61.8; DB 19; Length 795;
Best Local Similarity 48.7%; Pred. No. 1.6e-05;
Matches 168; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 25 gaagaacagataaagtcgcgcggaaatggaaagacagaagaagaatggcaag 84
DB 429 gaaggcagagagggcagagcagagagcagagagagggcagagcagagggg 488
QY 85 gagcattccagagagagaaggggaaatgatgacgtgtgtccttcacggacgcgttgag 144
DB 489 caggagcagagagggcagagcagagagcagagagggcagagagggcagagcagag 548
QY 145 cagttagaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 204
DB 549 ggagcagagagggcagagcagcagcagcagcagcagcagcagcagcagcagcagcag 608
QY 205 gagcagaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 264
DB 609 gagggcagagcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 668
QY 265 gaagtgctcacatcacccagacgatgagaggaactacagcgcaaaagtgaagccagc 324
DB 669 caggagagggcagagagggcagagagagagagagggcagagggcagagcagagcag 728
QY 325 cactggagtcgacccctcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 369
DB 729 aggtggagggcggcggtcgagagcagtgagggcggcggtcgag 773

RESULT 5

AAF22443
ID AAF22443 standard; cDNA; 968 BP.

XX AAF22443;

XX 26-MAR-2001 (first entry)

XX Human breast cancer associated antigen nucleotide sequence SEQ ID NO:22.

XX Human: breast cancer; gastric cancer; prostate cancer; diagnosis;

XX cancer associated antigen; cytostatic; cancer vaccine; ss.

XX Homo sapiens.

XX WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14749.

XX 28-MAY-1999; 99US-0136526.

XX 10-SEP-1999; 99US-0153454.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
XX antigen precursors, useful for diagnosing and treating a condition
XX characterized by expression of an abnormal amount of a protein, e.g.
XX cancer.

XX Claim 50; Page 278; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
XX represent nucleotide sequences encoding human breast, gastric and
XX prostate cancer associated antigen precursors (CAAP) respectively.
XX AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
XX represent human breast, gastric and prostate CAAP protein sequence
XX respectively. CAAPs have cytostatic activity and can be used in the
XX production of cancer vaccines. The human CAAP proteins, peptides, nucleic
XX acids or anti-CAAP antibodies are useful for diagnosing and treating a
XX condition characterised by expression of an abnormal amount of a protein,
XX e.g. cancer.

XX Sequence 968 BP; 219 A; 289 C; 329 G; 127 T; 4 other;

Query Match 12.7%; Score 59.6; DB 22; Length 968;
Best Local Similarity 66.2%; Pred. No. 5.3e-05;
Matches 86; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 117 cgtgtgctcttcagggacgcgttgagcagttagaagaagcggagcggagcgca 176
DB 164 ccttgacgacacatgaagcccgagcagagccctgtgtcagggagcggagcggagcgttcg 223
QY 177 gatcgggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 236
DB 224 caagggccagccctggagggagcagcggcttgacagcagcagcggcagcggagcagca 283
QY 237 gcggcgcaag 246
DB 284 gccgcgcgag 293

RESULT 6

AAAI2155
ID AAAI2155 standard; cDNA; 1618 BP.

XX AAAI2155;

XX 10-AUG-2000 (first entry)

XX Human X5L cDNA.

XX X5L; human; spermatogenesis protein; treatment; diagnosis; detection; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 113..1090

XX FT /*tag= a

XX FT /product= "spermatogenesis protein X5L"

XX DE19856882-C1.

XX 04-MAY-2000.

XX 10-DEC-1998; 98DE-1056882.

XX 10-DEC-1998; 98DE-1056882.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Sedlacek Z, Poustka A;

XX WPI; 2000-294180/26.

XX P-PSDB; AAY87646.

XX New spermatogenesis protein useful for investigation and modulation of
XX spermatogenesis, particularly for diagnosis and treatment of
XX spermatogenesis disorders -
XX Claim 3a; Fig 1; 14pp; German.

```
XX AC AAV3711;
XX 07-DEC-2000 (first entry)
XX DE Human PITSURE protein kinase isoform alpha2-2.
XX KW IRES; internal ribosome entry site; PITSURE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT misc_signal 907..1128
XX FT /*tag= a
XX FT /label= IRES
XX FT /note= "internal ribosome entry site"
XX PN W0200044896-A1.
XX PD 03-AUG-2000.
XX PE 26-JAN-2000; 2000MO-EP00643.
XX PR 26-JAN-1999; 99EP-0200216.
XX PA (VLAAR-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX PI Cornelis S, Beyaert R;
XX DR WPI; 2000-49931/44.
XX PT Nucleic acids encoding internal ribosome entry sequences useful for
XX PT directing protein expression in gene therapy procedures -
XX PS Disclosure; Page 32-33; 57pp; English.
XX CC The PITSURE protein kinase gene can be translated to give two isoforms,
XX CC p110 and p58. Transcription of p58 mRNA was found to start at an
XX CC internal ribosome entry site (IRES). The IRES element was found to be
XX CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX CC cycle. The present sequence is the human PITSURE protein kinase isoform
XX CC alpha2-2. This sequence contains the IRES. The IRES sequence and
XX CC fragments of it may be used to induce cell cycle dependent initiation
XX CC of translation in eukaryotic cells. Vectors containing the IRES may be
XX CC used for the preparation of compositions for the treatment of and/or
XX CC prevention of a disease by gene therapy. Such diseases may be cancer
XX CC and restenosis.
XX SO Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other;

Query Match 100.0%; Score 468; DB 21; Length 2471;
Best Local Similarity 100.0%; Pred. No. 2.5e-99;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caggaacgtcggaacgcgcacgcagaagaacagataagctcgccgagcggaagcg 60
DB 469 caggaacgtcggaacgcgcacgcagaagaacagataagctcgccgagcggaagcg 528
QY 61 cagaagaagaagaatagcaagagagcattccaggaagaagaaggggaatgagcgctg 120
DB 529 cagaagaagaagaatagcaagagagcattccaggaagaagaaggggaatgagcgctg 588
QY 121 tgcctctcaggagcgcttgagcagcttagaaaggaagcgagcgagcgcaagatg 180
DB 589 tgcctctcaggagcgcttgagcagcttagaaaggaagcgagcgagcgcaagatg 648
QY 181 cggagagcgaggaagcagcagcgaggaagaagcgcgagcgcgagcggaagcg 240
DB 649 cggagagcgaggaagcagcagcgaggaagaagcgcgagcgcgagcggaagcg 708
QY 241 cgcaagcgagcgagcgccgcagggagtgctgcacatccacgaaatgagagagcg 300
```

```
DB 709 cgcaagcgagcgagcgccgcagggagtgctgcacatccacgaaatgagagagcg 768
QY 301 tacagcgacaagaatgaagccagccactgagctgcagcccgctcgccgcgcggag 360
DB 769 tacagcgacaagaatgaagccagccactgagctgcagcccgctcgccgcgcggag 828
QY 361 cggctcgagctggagcagcgccggaagccagtaaaagaagaatgtgaaagaagc 420
DB 829 cggctcgagctggagcagcgccggaagccagtaaaagaagaatgtgaaagaagc 888
QY 421 ctgctgcgacttacagacatcagcgagcgagcgaggaagcaccagc 468
DB 889 ctgctgcgacttacagacatcagcgagcgagcgaggaagcaccagc 936

RESULT 4
AAV55830
ID AAV55830 standard; DNA; 795 BP.
XX AC AAV55830;
XX DT 18-NOV-1998 (first entry)
XX DE FLGA insert stabilising polypeptide encoding DNA.
XX KW Fusion protein: stabilising polypeptide; proteolytic degradation;
XX KW resistance; half-life; autoimmune disease; inflammation; nitro drug;
XX KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
XX KW nitroreductase protein; enzyme therapy; prodrg therapy; protease;
XX KW cancer; pathological condition; ss.
XX OS Epstein-barr virus.
XX FH Key Location/Qualifiers
XX FT CDS 1..788
XX FT /*tag= a
XX FT /product= "stabilising polypeptide"
XX PN W09822577-A1.
XX PD 28-MAY-1998.
XX PE 17-NOV-1997; 97WO-1B01508.
XX PR 25-JUN-1997; 97US-0048945.
XX PR 15-NOV-1996; 96US-0030986.
XX PA (MASU/) MASUCCI M G.
XX PI Masucci MG;
XX DR WPI: 1998-312463/27.
XX DR P-PSDB; AAM79128.
XX PT New fusion proteins resistant to proteolytic degradation -
XX PT comprising a core protein with a stabilising polypeptide comprising
XX PT a peptide sequence containing glycine repeats
XX PS Disclosure; Fig 3; 120pp; English.
XX CC This DNA encodes a stabilising polypeptide and is the FLGA insert of the
XX CC invention. The invention provides a method for increasing the resistance
XX CC of a core protein to proteolytic degradation that comprises linking or
XX CC inserting onto or into the core protein a stabilising polypeptide of
XX CC formula [(Glya)X(Glyb)Y(Glyc)Z]n where Glya, Glyb, Glyc are 1-6
XX CC sequential Gly residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met,
XX CC Phe, Pro or Thr and n can be anything between 1-66. X, Y and Z need not
XX CC be identical from n repeat to n repeat. Alternatively a nucleic acid
XX CC encoding the stabilising polypeptide can be linked onto or inserted into
XX CC a nucleic acid encoding a core protein. The fusion proteins of the
XX CC invention are more resistant to degradation by proteases and, thus, have
XX CC a longer half-life than the unfused core protein. The products can be
```


The PITRRE protein kinase gene can be translated to give two isoforms, p110 and p58. Transcription of p58 mRNA was found to start at an internal ribosome entry site (IRES). The IRES element was found to be cell cycle regulated, p58 is produced during the G2/M stage of the cell cycle. The present sequence contains a 5' fragment of the PITRRE protein kinase IRES and nucleotides upstream of the IRES. The IRES sequence and fragments of it may be used to induce cell cycle dependent initiation of translation in eukaryotic cells. Vectors containing the IRES may be used for the preparation of compositions for the treatment of and/or prevention of a disease by gene therapy. Such diseases may be cancer and restenosis.

50 Sequence 468 BP; 144 A; 100 C; 182 G; 42 T; 0 other;

Query Match	100.0%;	Score 468;	DB 21;	Length 468;
Best Local Similarity	100.0%;	Pred. No. 1.8e-99;		
Matches 468;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	caagaaacgctcgagaaacgcacatccagaaagaaacagagataaagcttcgcccgaggaaatcvgagaaaga	60
Db	1	caagaaacgctcgagaaacgcacatccagaaagaaacagagataaagcttcgcccgaggaaatcvgagaaaga	60
QY	61	caagaaagaaagagaaatctgacaaagagacatctccagagagagaaaggggagatgacgtg	120
Db	61	caagaaagaaagagaaatctgacaaagagacatctccagagagagaaaggggagatgacgtg	120
QY	121	tgacctcttcagagagacgctctgagacagctaaagaaagaaacgagagagcgagacgaagatg	180
Db	121	tgacctcttcagagagacgctctgagacagctaaagaaagaaacgagagagcgagacgaagatg	180
QY	181	cagagacacagacaagaaagagacagacgagagacaaagaaagagacgagacgagcvgagcagagagacgag	240
Db	181	cagagagacagacaagaaagagacagagcagcvgagaaagagagacgagacgagcvgagcagagagacgag	240
QY	241	cagcaagagcagagagcccgacagagaaagctgtccgcacatccacacgaacgaatcgaagaaagac	300
Db	241	cagcaagagcagagagcccgacagagaaagctgtccgcacatccacacgaacgaatcgaagaaagac	300
QY	301	tacaaacgaaacaaagctgaagagccagacaaatctgagatcctgaagcccgcttcgagccgacgagag	360
Db	301	tacaaacgaaacaaagctgaagagccagacaaatctgagatcctgaagcccgcttcgagccgacgagag	360
QY	361	cagcttcgaagctcvgagaaacgagccgagaaagccagctaaagaaagagaaatctgaagaaagagac	420
Db	361	cagcttcgaagctcvgagaaacgagccgagaaagccagctaaagaaagagaaatctgaagaaagagac	420
QY	421	cagcagctcgaactacagagacatcagagaaagacgagaaagaaagaaacacagc	468
Db	421	cagcagctcgaactacagagaaatcagagaaacgagaaagaaagaaagaaacacagc	468

RESULT	2
AAA73714	
ID	AAA73714 standard; DNA; 660 BP.

AC AAA73714;

DT 07-DEC-2000 (first entry)

Human PITSLRE protein kinase IRES and upstream sequence.

KW
IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
cancer; rectal cancer; 258-310

KW cancer; restenosis; p58; p110; protein kinase; ds.
XY

OS Homo sapiens.

	Key	Location/Qualifiers
FH		
FT	misc_signal	439..660

```
FT /label= IRES
FT /note= "Internal ribosome entry site"
```

PN W0200044896-A1

PD 03-AUG-2000

PF 26-JAN-2000; 2000WO-EP00643.

PR 26-JAN-1999; 99EP-0200216.
XX

PA (VLA-) VLAMS INTERUNIVERSITAIR INST BIOTECHNOG.
VY

PI Cornelis S, Beyaert R;
v.v.

DR WPI; 2000-499331/44.
YY

PT Nucleic acids encoding internal ribosome entry sequences useful for directing protein synthesis in cells

XX Claim 10: Page 34. Error: Estimating
PS

CC The PITSRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence contains the PITSRE protein kinase
CC IRES and upstream nucleotides. The IRES sequence and fragments of it
CC may be used to induce cell cycle dependent initiation of translation in
CC eukaryotic cells. Vectors containing the IRES may be used for the
CC preparation of compositions for the treatment of and/or prevention of
CC a disease by gene therapy. Such diseases may be cancer and restenosis
XX
XX Sequence 660 BP; 210 A; 128 C; 261 G; 61 T; 0 other;

Sequence 660 BP; 210 A; 128 C; 261 G; 61 T; 0 other;

Query Match	100.0%;	Score 468;	DB 21;	Length 660;
Best Local Similarity	100.0%;	Pred. No. 1.9e-99;		
Matches 468; Conservative	0;	Mismatches	0;	Indels 0;
		Gaps	0;	

QY 1 caccgaacgtcggaaacgacatcgagaagaacagatgaataagctcgccgggaatgggaaaga 60

Db 1 caccgaacgtcggaaacgacatcgagaagaacagatataagctcgccgggaatggaaga 60

QY 61 cagaagagaggaatgccaagggagcatccagagagagaagggggaatgatggcgtg 12

Db 61 cagaagagaaggaatgcaagggagcatccagagagaaaggggaaatgatgcgctg 12

121 tgcctcttcaggaccgcttgagcagttagaagaagcgggagcgggagcgcgaagtg 18

Db 121 tgcctctcagggaccgcttgagcagttagaagaagcgggagcggcaagatg 18

181. *cgggagcagcagaaaggagcaggagcgcgagcggagagcgg* 24

101 cgggagcagcaggagcgggagcaggagcgggagcgggagcgggagcggg 24

[illegible]

30

[illegible]

----- ၁၀

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

— — — — —

RESULT	3
AAA73711	
ID	AAA73711 standard; DNA; 2471 BP.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 19:03:36 : Search time 355.05 Seconds
(without alignments)
1130.061 Million cell updates/sec

Title: US-09-915-060-5

Perfect score: 468

Sequence: 1 cacgaacgtcggaacgaca.....acagcagaggaagaccacgc 468

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_1101:*

1: /SIDS2/gcgdata/geneseq/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468	100.0	468	21	AAAT3713
2	468	100.0	660	21	AAAT3714
3	468	100.0	2471	19	AAAT3711
4	61.8	13.2	795	19	AAV55830
5	59.6	12.7	968	22	AAAT2443
6	59.6	12.7	1618	21	AAAT12155
7	59.6	12.7	1638	21	AAAT6636
8	59.6	12.7	1645	22	AAAT58915
9	59.6	12.7	1648	22	AAAT60701
10	59.6	12.7	2875	21	AAAT12156
11	59.4	12.7	2800	22	AAAT07131

12	59.4	12.7	2805	22	AAAT07130
13	58	12.4	12001	16	AAAT07213
14	57.8	12.4	2803	22	AAAT07132
15	57.6	12.3	799	19	AAV55831
16	57.6	12.3	1926	21	AAAT50254
17	57.6	12.3	1926	22	AAAT82902
18	57.6	12.3	2580	21	AAAT5454
19	57.6	12.3	5452	20	AAAT0923
20	57.6	12.3	8705	20	AAAT23778
21	57.6	12.3	9600	19	AAAT1683
22	57.6	12.3	10380	20	AAAT22248
23	57.6	12.3	10596	14	AAAT51731
24	57.6	12.3	10596	17	AAAT40348
25	57.6	12.3	10596	20	AAAT15650
26	57.6	12.3	16080	21	AAAT59553
27	57.2	12.2	9551	20	AAAT22301
28	56.2	12.0	1925	20	AAAT90924
29	56	12.0	699	21	AAAT36119
30	55.6	11.9	16442	18	AAAT83006
31	54	11.5	51259	18	AAAT83007
32	53.8	11.5	2301	19	AAAT20445
33	53.8	11.5	2301	21	AAAT09303
34	53.4	11.4	696	21	AAAT48939
35	53.2	11.4	397	20	AAAT89891
36	51.6	11.0	49999	20	AAAT23895
37	51.4	11.0	168	16	AAAT84833
38	51.4	11.0	171	16	AAAT84834
39	51.4	11.0	195	16	AAAT84831
40	51.4	11.0	234	16	AAAT84832
41	51.2	10.9	10732	21	AAAT10594
42	51	10.9	465	22	AAAT25521
43	51	10.9	465	22	AAAT15993
44	51	10.9	486	22	AAAT16341
45	51	10.9	486	22	AAAT38923

ALIGNMENTS

RESULT 1

AAAT3713

AAAT3713 standard. DNA; 468 BP.

AAAT3713;

07-DEC-2000 (first entry)

5' extension of PITSURE protein kinase internal ribosome entry site.

IRES; internal ribosome entry site; PITSURE; human; gene therapy;

KW cancer; retinoblastoma; p53; p110; protein kinase; ds.

XX Homo sapiens.

XX WO200044896-A1.

XX 03-AUG-2000.

XX 26-JAN-2000; 2000MO-EP00643.

XX 26-JAN-1999; 99EP-0200216.

XX (VIAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Cornelis S, Beyaert R;

XX WPI; 2000-499331/44.

XX Nucleic acids encoding internal ribosome entry sequences useful for

XX directing protein expression in gene therapy procedures

XX Claim 9; Page 33; 57pp; English.

Wed Dec 5 08:45:08 2001

us-09-915-060-5.rge

Page 11

Db 909 CTGCTGTCCGACTTACAGACATCAGCGACGACGAGAGAGACCAGC 956

Search completed: December 4, 2001, 18:54:37
Job time: 9173 sec


```
|||||
Db      829  CGGTTCCAGTGGAGAGCGCCGGAACCCAGTAAAGAAAGAAAATGGAAGAGGAGC 888
QY      421  ctgctgtccgacttacaggacatcagcagcaggaaggaagccagc 468
          |||||||
Db      889  CTGCTGTCCGACTTACAGACATCAGACAGCAGAGGAAGACCAGC 936

RESULT  13
LOCUS   HSU04824      2477 bp      mRNA      PRI      08-JUL-1994
DEFINITION Human protein kinase PITSLRE alpha 2-1 mRNA, complete cds.
ACCESSION U04824
VERSION   U04824.1 GI:507167
KEYWORDS
SOURCE    human.
ORGANISM Homo sapiens
            Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2477)
AUTHORS   Xiang,J., Lahti,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE     Molecular cloning and expression of alternatively spliced PITSLRE
          Protein kinase isoforms
JOURNAL   J. Biol. Chem. 269, 15786-15794 (1994)
MEDLINE   94253170
REFERENCE 2 (bases 1 to 2477)
AUTHORS   Kidd,V.J.
TITLE     Direct Submission
          Submitted (04-JAN-1994) Vincent J. Kidd, St. Jude Children's
          Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
          Memphis, TN 38101, USA
FEATURES
  source
    1..2477
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /map="1p36-2"
    /cell_line="Hela S3"
    /cell_type="epitheloid"
    /tissue_type="cervix"
    /clone_lib="Hela cDNA library"
    /dev_stage="adult"
    1..2477
    /codon_start=1
    112..2451
    /product="PITSLRE alpha 2-1"
    /protein_id="AA19586.1"
    /db_xref="GI:507168"
    /translation="MGDEKDSMKVKTLDLLEIQRKRRKDEEKAELIKRLNSDDRDK
    /RSLEEGELRDHMEITIRNSPYRREDSMEDRGEEDSLAIKPPQMSRKEVHHRKD
    EKREKRHRSHSAEGCKHARVKKRERERRRHREDDKAREREROKRREMAREH
    SRERRLRLQLEKRRERERERERERERERERERERERERERERERERERERERER
    REDYSDKVAASHMSRSPRRPRRFLQGRKPKYKRNKREBDLSDLDIDISERKT
    SSAFESSAEGSGSEEEEEESEEESESESESESESESESESESESESESESESESE
    EYSEEMSEDEERENHLLVYPSERFDSDGSESESESESESESESESESESESESESE
    DSPALSPILQELQELPKYLPALQCRSVEEFQCNRIEEGTGVVYAKAKDEIVALL
    KRKMEKEKGGPFTLSREINTILKAQHNITVREIVGSSNMNDKIYIMANYVEDLK
    SLMEYMKQPLPEGVVTLMIOLLRGYKHLHDMMIILHRLDKTNSILSHGILKVGDFG
    LAREYSPILAKATPVVYTLMYRRLPRLLELLAKEXSTRVDMMSVGCITGELLTKPLPFG
    KSEIDQINKVFKDGLTPSEKIMPGYSELPAVKMFTSRHPYNNLRKRFALLSDGCFD
    LMKKFLTYPPGRIRISADGLKHEYRETPLPIDPSMFPYPAKSEQORVKNRGTSPRP
    EGGIGYSQGDGDLKETGPHLTTTNOGASAGPFSLKF"

BASE COUNT  701 a      603 c      811 g      362 t

ORIGIN
```

Query Match 86.3%; Score 404; DB 9; Length 2477;
Best Local Similarity 94.2%; Pred. No. 2,1e-69;
Matches 441; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

```
QY      1  caagaaactcgaaagacatcgagaagacagataaagctcgccggaatggaaga 60
          |||||||
Db      502  CAGGAACGTCGGAAGACATCGAAGAACAGATMAACCTCCGCGGAATGGGAAGA 561
```

```
QY      61  cagaagagaaggaaatgccaagggagcatctccagaagagaagaagggaatgagcgtg 120
          |||||||
Db      562  CAGAAAGAGGGAATGCGCAAGGGAGCATCTCCAGAGAGCA----- 603
QY      121  tgcctctcaaggaccgcttggagcagltagaagaagcggagcggagcgcaagatg 180
          |||||||
Db      604  -----AGGACCGCGCTTGGAGCAGTTAGAAAGAAAGCGGACCGGACCGCAAGATG 654
QY      181  cggagacagcagaagagcaggcagggaagagcgagcggcggcggaagagcg 240
          |||||||
Db      655  CGGGACACACAGAGAGAGCAGCGGAGCAGAGAGCGGACCGCGGAGGAGAGCGG 714
QY      241  ccgaagagcggagagcccgccaggggaagtgtctgcacatccagcaagatagaagagc 300
          |||||||
Db      715  CGCAAGGACGGGAGCGCCCGCAGGGAGAGTGTGCAACATCACCGAATGATGAGAGGAC 774
QY      301  taacagcgaacagtgaagaagccagccactgagctgcgaagccgcctcggccgcyggag 360
          |||||||
Db      775  TACAGCGACAAAGTGAAAGCCAGCCACTGTGAGTCCGACGCCCTCGGCCCGCGGAG 834
QY      361  cggcttgagattggagagcggccggagagccagtgtaaaagaagagaagaatggaagagac 420
          |||||||
Db      835  CGGTTGAGATTGGGAGACCGCGCGGAGCCAGTAAAGAAAGAGCAAAATGGAAGAAAGGAGC 894
QY      421  ctgctgtccgacttacaggacatcagcagcagcagcaggaaggaagaccagc 468
          |||||||
Db      895  CTGCTGTCCGACTTACAGACATCAGCCAGCAGAGAGAGAACCCAGC 942
```

```
RESULT  14
LOCUS   AF067512      2486 bp      mRNA      PRI      07-NOV-1998
DEFINITION Homo sapiens PITSLRE protein kinase alpha SV1 isoform (CDC2L1)
ACCESSION AF067512
VERSION   AF067512.1 GI:3850303
KEYWORDS
SOURCE    human.
ORGANISM Homo sapiens
            Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2486)
AUTHORS   Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
          Ambros,P.F. and Kidd,V.J.
TITLE     Duplication of a genomic region containing the Cdc2L1-2 and
          MPE21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2
          Genome Res. 8 (9), 929-939 (1998)
JOURNAL   98424414
MEDLINE   98424414
REFERENCE 2 (bases 1 to 2486)
AUTHORS   Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
          Ambros,P. and Kidd,V.J.
TITLE     Direct Submission
          Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
          Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES
  source
    1..2486
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="1"
    /map="1p36.3"
    1..2486
    /gene="CDC2L1"
    112..2460
    /gene="CDC2L1"
    /note="corresponding genomic sequence deposited as GenBank
    Accession Numbers AF080678-AF080688, AF092429, and
    AF092430"
    /codon_start=1
    /product="PITSLRE protein kinase alpha SV1 isoform"
    /protein_id="AAC72077.1"
    /db_xref="GI:3850304"
    /translation="MGDEKDSMKVKTLDLLEIQRKRRKDEEKAELIKRLNSDDRDK
    RDSLEEGELRDHMEITIRNSPYRREDSMEDRGEEDSLAIKPPQMSRKEVHHRKD
```


Query Match	86.3%;	Score 404;	DB 9;	Length 2439;
Best Local Similarity	94.2%;	Pred. No. 2.1e-69;		
Matches 441; Conservative	0;	Mismatches 0;	Indels 27;	Gaps 1;

HSU04817	11	RESULT	11
LOCUS			
DEFINITION	HSU04817	2444 bp	mRNA
ACCESSION	U04817		PRI
VERSION	U04817.1	GI:507161	08-JUL-1994
KEYWORDS			complete cds.
SOURCE			
ORGANISM			
	human.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 2444)		
AUTHORS	Xiang,J., Iatli,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.		
TITLE	Molecular cloning and expression of alternatively spliced PITSURE protein kinase isoforms		
JOURNAL	J. Biol. Chem. 269, 15786-15794 (1994)		
MEDLINE	94253170		
REFERENCE	2 (bases 1 to 2444)		
AUTHORS	Kidd,V.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St., Memphis, TN 38101, USA		
FEATURES			
source	location/Qualifiers		
	1..2444		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/map="lp36-2"		
	/cell_line="Hela S3"		
	/cell_type="epitheloid"		
	/tissue_type="cervix"		
	/clone_lib="Hela CDNA library"		
	/dev_stage="adult"		
	1..2444		
mRNA	112..2418		
cds	/codon_start=1		
	/product="PITSURE alpha 2-3"		
	/protein_id="AA19583.1"		
	/db_xref="GI:507162"		
	/translation="MGDEKSDMKVYTLDEILLOEKRRKEOEKAEIKRLKNSDPRDSK		

[illegible]

BASE COUNT	690 a	591 c	789 g	366 t	
ORIGIN					
Query Match	99.3%	Score 464.8;	DB 9;	Length 2436;	
Best Local Similarity	99.6%;	Pred. No. 3e-81;			
Matches 466;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
OY	1	caacgaacgtcggaaacacgacatcagagaaacaggaataaagctccgcgcgggaatgggaaaag 60			
DB	434	CACGCAACGTCGGAAACGACATCGAAGAAGACAGGATTAAGCTCGCGGGGAATGGGAAAGA 493			
OY	61	cagaagagaaaggaatgycgaaggagcattccagagagagaaagggggaatgatatgcctg 120			
DB	494	CAGAAAGAGAGGGAATGGCAGAGGAGCATTCACAGAGAGAAAGGGGAGCATGATGCTTT 553			
OY	121	tgcctcttcgaaggacgcgcttggagcagttagaagaacgcygagcggagcgcgaagtgc 180			
DB	554	TCCCTCTTCAGGGACCGCTTGGAGCAGTTAGAAAGAAAGCGGAGCGGAGCGCAAGATG 613			
OY	181	cggagagcagcaagaagagcagcagcgaggaagaaggagcgcgagcggcgcggaagacg 240			
DB	614	CGGGAGCAGCAAAAGAGAGCGCGGAGCAGAGAGAGCCGCGCGCGGCGGAGAGCGG 673			
OY	241	cgcgaagagcgaggaagccgcgagggaaatgctctgacataccagaaacatgagagagac 300			
DB	674	CGCAAGAGAGCGGAGAGCCGCGAGAGTGTGTGCATCATCCGAAACATAGAGAGAGAC 733			
OY	301	tacagcgacaaagtgaagccagccacttggagtcgcaagccgcctcggcgccgcggagag 360			
DB	734	TACACGCAAAAGTGAAGCCAGCCACTGAGATCTGCACACCCGCTCGCGCCGCGGAG 793			
OY	361	cggctcgcagtttgaggagacgcgcggagagccagtaaaagaagaagaataatgaagaaggac 420			
DB	794	CGGTCGAGTTGGGAGAGCGCCGGAAGCCAGATGAAGAGCAAAATGGAAGAAAGGAGC 853			
OY	421	ctgcgtccgcactacagagacatccgcagacagcgaagagaaacccagc 468			
DB	854	CGGCTCGCATTTACAGGACATCTACGACACAGCAGAGAGAGAGACACGAC 901			
RESULT	8				
LOCUS	AF067518	2323 bp	mRNA	PRI	07-NOV-1998
DEFINITION	Homo sapiens PIRSLR protein kinase beta SV8 isoform (CDC2L2) mRNA, complete cds.				
ACCESSION	AF067518				
VERSION	AF067518.1 GI:3850315				
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 2323)				
AUTHORS	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P.F. and Kidd,V.J.				
TITLE	Duplication of a genomic region containing the Cdc2L1-2 and Mh2L1-22 genes on human chromosome 1p36.3 and their linkage to D1Z2				
JOURNAL	Genome Res. 8 (9), 929-939 (1998)				
MEDLINE	98424414				
REFERENCE	2 (bases 1 to 2323)				
AUTHORS	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA				
FEATURES	Location/Qualifiers				
SOURCE	1..2323				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="1"				

Query Match 100.0%; Score 468; DB 6; Length 468;
 Best Local Similarity 100.0%; Pred. No. 9,3e-82;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BASE COUNT 144 a 100 c 182 g 42 t
 ORIGIN /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1 cagaacgtcgaagacacatcgagaagacagatataagctccgggaattggaaaga 60
 |||||||
 1 CAGAACGTCGGAACGACATCGAAGAACAGATTAAGCTCCGCGAATGGGAAGA 60
 |||||||
 61 cagaagaagaggaaatgcaagaggacattccagaagagaaggggaattgagctg 120
 |||||||
 61 CAGAAGAGAGGAATGCAAGGAGCATTCAGAGAGAAAGGGGAATGATGGCGTG 120
 |||||||
 121 tgccttcaaggagccgcttgtagcagttagaagaagacgggagcggcaagatg 180
 |||||||
 121 TGCTCTTCAAGGACCGCTTGAGCATTAAGAAAGCAAGCGGAGCGCAAGATG 180
 |||||||
 181 cggagcagcagaagaagcagcggagcagaagcggcggcggcggcggcggcggc 240
 |||||||
 181 CGGAGCAGCAGAGAGAGCAGCGGAGCAGAAAGAGCGCGCGCGAGAGCGG 240
 |||||||
 241 cgcgaagagcgggagagccgcgaggaagtgtctgcacatcagcagatgagagagac 300
 |||||||
 241 CGCAAGGACGGGAGGCGCGCAGGGAGAGTGTGCACATCAGCAACGATGAGAGGAGC 300
 |||||||
 301 taagagacaagaagtgaagccagccacttgagtcgagccgctcgccgcggcgagag 360
 |||||||
 301 TACAGCGACAAAGTGAAGCCAGCCACTGAGTCCGAGCCCGCTCGCGCGCGGAG 360
 |||||||
 361 cggctcgaagltgggaagcggcggagccagtaaaagaagaagaatggaagaagagac 420
 |||||||
 361 CGGTCGAGTGGGAGAGCGCGGAGCCAGTAAGAAAGAAATGAGAGAAAGGAGC 420
 |||||||
 421 ctgctgtccgacttacagagacatcagcgaagcgaaggaagaccagc 468
 |||||||
 421 CTGCTGTCCGACTTACAGGACATCAGCGACGAGAGAGAGACCGAGC 468
 |||||||

RESULT 2
 AX033425 660 bp DNA PAT 21-SEP-2000
 LOCUS AX033425
 DEFINITION Sequence 6 from Patent WO0044896.
 ACCESSION AX033425
 VERSION AX033425.1 GI:10280186
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 660)
 AUTHORS Beyerle, R. and Cornelis, S.
 TITLE Internal ribosome entry site (IRES), vector containing same and
 uses thereof
 JOURNAL Patent: WO 0044896-A 6 03-AUG-2000;
 VILAMS INTERNIV INST BIOTECH (BE); BEYAERT RUDI (BE); CORNELIS
 SIGRID (BE)
 FEATURES
 source 1..660
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 210 a 128 c 261 g 61 t
 ORIGIN

Query Match 100.0%; Score 468; DB 6; Length 660;
 Best Local Similarity 100.0%; Pred. No. 8,8e-82;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagaacgtcgaagacacatcgagaagacagatataagctccgggaattggaaaga 60
 |||||||
 DB 1 CAGAACGTCGGAACGACATCGAAGAACAGATTAAGCTCCGCGAATGGGAAGA 60
 |||||||
 QY 61 cagaagaagaggaaatgcaagaggacattccagaagagaaggggaattgagctg 120
 |||||||
 DB 61 CAGAAGAGAGGAATGCAAGGAGCATTCAGAGAGAAAGGGGAATGATGGCGTG 120
 |||||||
 QY 121 tgccttcaaggagccgcttgtagcagttagaagaagacgggagcggcaagatg 180
 |||||||
 DB 121 TGCTCTTCAAGGACCGCTTGAGCATTAAGAAAGCAAGCGGAGCGCAAGATG 180
 |||||||
 QY 181 cggagcagcagaagaagcagcggagcagaagcggcggcggcggcggcggcggc 240
 |||||||
 DB 181 CGGAGCAGCAGAGAGAGCAGCGGAGCAGAAAGAGCGCGCGCGAGAGCGG 240
 |||||||
 QY 241 cgcgaagagcgggagagccgcgaggaagtgtctgcacatcagcagatgagagagac 300
 |||||||
 DB 241 CGCAAGGACGGGAGGCGCGCAGGGAGAGTGTCTGCACATCAGCAACGATGAGAGGAGC 300
 |||||||
 QY 301 taagagacaagaagtgaagccagccacttgagtcgagccgctcgccgcggcgagag 360
 |||||||
 DB 301 TACAGCGACAAAGTGAAGCCAGCCACTGAGTCCGAGCCCGCTCGCGCGCGGAG 360
 |||||||
 QY 361 cggctcgaagltgggaagcggcggagccagtaaaagaagaagaatggaagaagagac 420
 |||||||
 DB 361 CGGTCGAGTGGGAGAGCGCGGAGCCAGTAAGAAAGAAATGAGAGAAAGGAGC 420
 |||||||
 QY 421 ctgctgtccgacttacagagacatcagcgaagcgaaggaagaccagc 468
 |||||||
 DB 421 CTGCTGTCCGACTTACAGGACATCAGCGACGAGAGAGAGACCGAGC 468
 |||||||

RESULT 3
 HSU07705 2465 bp mRNA PRI 08-JUL-1994
 LOCUS HSU07705
 DEFINITION Human protein kinase PITSLRE isoform PBETA22 (p58CDC2L1) mRNA,
 complete cds.
 ACCESSION U07705
 VERSION U07705.1 GI:507428
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2465)
 AUTHORS Xiang, J., Lahti, J.M., Grenet, J.A., Easton, J.B. and Kidd, V.J.
 TITLE Molecular cloning and expression of alternatively spliced PITSLRE
 protein kinase isoforms
 JOURNAL J. Biol. Chem. 269, 15786-15794 (1994)
 MEDLINE 94253170
 REFERENCE 2 (bases 1 to 2465)
 AUTHORS Kidd, V.J.
 TITLE Direct Submission
 JOURNAL Submitted (14-MAR-1994) Vincent J. Kidd, St. Jude Children's
 Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
 Memphis, TN 38101, USA
 FEATURES
 source 1..2465
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1p36-2"
 /cell_line="HeLa S3"
 /cell_type="epitheloid"
 /tissue_type="cervix"
 /clone_id="HeLa CDNA library"
 /dev_stage="adult"
 1..2465
 112..2439
 /gene="p58CDC2L1"
 112..2439
 /gene="p58CDC2L1"

mRNA
 gene
 CDS
 /gene="p58CDC2L1"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:54:31 ; Search time 2399.05 seconds
(without alignments)
3218.224 Million cell updates/sec

Title: US-09-915-060-5

Perfect score: 468
Sequence: 1 caccagactcggaagcagaca.....acagcagagagagaccagc 468

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vi: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rod: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468	100.0	468	6	AX033424
2	468	100.0	660	6	AX033425
3	468	100.0	2465	6	HSU07705
4	468	100.0	2471	6	AX033422
5	468	100.0	2471	6	HSU04816
6	464.8	99.3	2329	9	AF067525
7	464.8	99.3	2436	9	AF067520
8	404	86.3	2323	9	AF067518
9	404	86.3	2409	9	AF067521
10	404	86.3	2439	9	AF067519
11	404	86.3	2444	9	HSU04817
12	404	86.3	2471	9	HSU07704
13	404	86.3	2477	9	HSU04824
14	404	86.3	2486	9	AF067512
15	404	86.3	2500	9	AF067514
16	404	86.3	2533	9	AF067516
17	355	75.9	2362	9	AF067523
18	355	75.9	2448	9	AF067522
19	355	75.9	2525	9	AF067515
20	348.6	74.5	2544	9	AK000081
21	296.8	63.4	3161	10	MUSCDEPK
22	203.4	43.5	2340	9	HSU04818
23	203.4	43.5	2349	9	AF067517
24	181	38.7	1715	9	AF174497
25	166	35.5	110608	9	HS28383
26	165	34.6	297	9	CDC2LIS06
27	161.8	34.6	297	9	CDC2LIS06
28	151	32.3	1936	10	MUSP58GTA
29	138.4	29.6	712	9	AF067524
30	127.4	27.2	532	9	CDC2LIS07
31	127.4	27.2	533	9	CDC2LIS07
32	105.4	22.5	539	9	CDC2LIS05
33	105.4	22.5	761	9	CDC2LIS05
34	105	22.4	2226	9	AF067513
35	100	21.4	551	9	AF067527
36	100	21.4	700	9	AF067529
37	99.4	21.2	7218	6	AF6494
38	98	20.9	579	9	AF067526
39	98	20.9	728	9	AF067528
40	79	16.9	480	9	CDC2LIS08
41	79	16.9	592	9	CDC2LIS08
42	72.4	15.5	484	5	CPTSLRE02
43	67.2	14.4	178022	2	AC018348
44	67.2	14.4	179880	2	AC091304
45	66.2	14.1	238688	2	AC079818

ALIGNMENTS

RESULT 1
AX033424
LOCUS AX033424 468 bp DNA
DEFINITION Sequence 5 from Patent WO0044896.
ACCESSION AX033424
VERSION AX033424.1 GI:10280185
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 468)
AUTHORS Beyaert, R. and Cornelis, S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
JOURNAL Patent: WO 0044896-A 5 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)
FEATURES
source location/Qualifiers
1..468


```

Db      334 GCAGATATGAGCTCAGGTTCTGGAAGAGGAGGACGAGGAACGAGAAACGAGGAAGACGA 393
Oy      538 gagagagaagaggagaccagctgaagaatcagagagagaagaagaaga-----584
Db      394 ACGAAGAACGGGAGACACCATGTAACGAACCTCAGAGGAAGAACGAGGAAGAAAGAGA 453
Oy      585 -----gagagagagagagacggcagcaactctggagagagcatcagagagctgcgaag 640
Db      454 GGACGGAGGAAGAGGAGAGACTGGAGCAACTCTGAGAGAGAGCCTCTGCACAGTCCGACAGA 513
Oy      641 aagtaagtgaggaagaatg 660
Db      514 AAGTCAGTCAGTATGAGGAATG 533

RESULT 14
BG943585
LOCUS      BG943585      326 bp      mRNA      EST      11-JUN-2001
DEFINITION ax39f04.x1 Proliferating Human Erythroid Cells (LCB:ax library)
ACCESSION BG943585
VERSION    BG943585
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 326)
            Gulin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.
            Gene expression in proliferating human erythroid cells
            Genomics 59 (2), 168-177 (1999)
COMMENT    Contact: Jeffery L. Miller
            Laboratory of Chemical Biology
            National Institute of Diabetes and Digestive and Kidney Diseases
            Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
            20892, USA
            Tel: 301 402 2373
            Fax: 301 435 5148
            Email: jml7@nih.gov
            DNA Sequencing and analyses by National Institutes of Health
            Intramural Sequencing Center (NISC).
            Plate: 39 row: f column: 04
            Seq primer: 21M13 forward primer (ABI).
FEATURES
     source
         1..326
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /clone="ax39f04"
         /clone_id="Proliferating Human Erythroid Cells (LCB:ax library)"
         /sex="unknown"
         /tissue_type="Blood"
         /cell_type="Erythroid Cells"
         /cell_line="Primary Culture of Peripheral Blood Mononuclear Cells"
         /dev_host="SOLR"
         /note="Organ: blood, Vector: Lambda ZAP II, Site_1: EcoRI; Site_2: EcoRI; 65,000 proliferating erythroid cells from the buffy coat of a blood donation were obtained by flow cytometric separation after a 5-day culture period in the presence of erythropoietin. Total RNA was purified from the sorted cell population using TRIzol reagent. RNA (0.3 ug) was converted into double stranded cDNA using Clontech's Capfinder cDNA library Construction Kit (Clontech) according to the manufacturer's protocol and cloned into EcoRI digested Lambda Zap II vector (Stratagene). The phage library was amplified once prior to in vivo excision in SOLR cells. Individual colonies were grown, and the cDNA inserts were sequenced in high throughput (NIH Intramural sequencing center http://www.nisc.nih.gov/)."

```

BASE COUNT	109	a	67	c	128	g	22	t
ORIGIN								
Query Match	41.2%;	Score	271.6;	DB	11;	Length	326;	
Best Local Similarity	96.9%;	Pred.	No. 2.3e-35;					
Matches	277;	Conservative	0;	Mismatches	9;	Indels	0;	Gaps
OY	130	aggagaccgccttgaagcaagttaagaagaagcgggagcgggagcgaatgtccggagacag	189					
Db	35	AGGACCGCTTGAAGCA GTTAGAAAGAAAGCAGCGGAGCCGAGACGCAGAATCCGGAGACAG	94					
OY	190	caagaagaagcagcgggagcagaagagcgccgagcgcgcgcgagagcgcgcaaaagag	249					
Db	95	CAGAAGGAGCAGCGGGGAGCAGAAGGAGGCCGAGCGCGGGGCGAGACGCGCGAGAGAG	154					
OY	250	cgggagggcccgcagggagaagtgtctgcatacatcacgaaagatgagagagactaacgacg	309					
Db	155	CGGAGAGCCCCGACAGGGAAGTGTCGTGCACATCACGACAACGACTAGAGAGACTACAGCGAC	214					
OY	310	aaaatgaagccaagccaccttgagttgcacaaccccccttcgacgcgcgcgagagcgttcag	369					
Db	215	AAAATGAAAGCCACACCCTGTGATCTCCACGCCGCCCTCGGGCCGCCGCGGAGACGGTTGAG	274					
OY	370	ttggagacgcgccggaagccagtaaaaaagaagaanaaatyygaagaa	415					
Db	275	TTCGGAGACGCGCGAAGCCAGAAAAA AAAAAAAAAAAAAAAAAAAAAA	320					
RESULT	15							
BG969359	758	bp	mRNA	EST	12-JUN-2001			
LOCUS	60283684.F1	NCL_CGAP_Kid14	Mus musculus	cdna clone IMAGE:4971155				
DEFINITION	5', mRNA sequence.							
ACCESSION	BG969359							
VERSION	BG969359.1	GI:14356996						
KEYWORDS	EST.							
SOURCE	house mouse.							
ORGANISM	Mus musculus							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
TITLE	1 (Phases 1 to 758)							
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.							
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LHAM10956 row: a column: 12 High quality sequence start: 3 High quality sequence stop: 731. Location/Qualifiers 1..758							
FEATURES	source							
	/organism="Mus musculus"							
	/strain="FVB/N"							
	/db_xref="taxon:10090"							
	/clone="IMAGE:4971155"							
	/clone_1ib="NCL_CGAP_Kid14"							
	/lab_host="DH10B (TI phage-resistant)"							
	/note="Organ: kidney; Vector: pCMV-Sport6; Site_1: NCI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI-CGAP Library."							
BASE COUNT	254	a	144	c	254	g	106	t
ORIGIN								


```
/clone_lib="NIH_MGC_87"  
/tissue_type="mammary adenocarcinoma, cell line"  
/lab_host="DH10B (phage-resistant)"  
/note="Organ: breast; Vector: PCMV-Spore6; Site: 1: NotI;  
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.383 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH_MGC Library."  
BASE COUNT      272 a      165 c      270 g      97 t  
ORIGIN  
  
Query Match      47.6%; Score 314.2; DB 11; Length 804;  
Best Local Similarity 84.5%; Pred. No. 2.4e-42;  
Matches 388; Conservative 0; Mismatches 43; Indels 28; Gaps 2;  
  
QY      1  cagaacgttcggaagcaatcgaagaagcaaggaatgaagccgcgcgggaatgggaaga 60  
      |||||||  
Db      358  CACGAACGTGGAAACGCGATCGAGAAGACAGATAAAGCTCGCGGGAATGGAAAGA 417  
  
QY      61  cagaagagaaggaatggcaagagacatccagaagagaaggggaatgatgctg 120  
      |||||||  
Db      418  CACAAGAGAAGGAGATGGCAAGGAGCATTCACAGAGAGAA----- 459  
  
QY      121  tgcctctcagggaccgcttgagacgttaagaagaagcggagcggagcgcgaatg 180  
      |||||||  
Db      460  -----AGGGACCGCTTGGAGCAGTTAGAAAGGAAAGCGGAGCGCAGCAAGATG 510  
  
QY      181  cggagacagcagaagagcggagcggagcgaagaagcggcggcggcggagagcg 240  
      |||||||  
Db      511  CGGAGACAGAGAGAGAGCGCGGAGCAGAAAGAGAGCGGAGCGGCGGAGAGCGG 570  
  
QY      241  cgaagagcggagagcccgcaaggaagtgtctgcacatcacgaaacgatagaagagc 300  
      |||||||  
Db      571  CGCAAGAGAGGAGGCGCCGACAGCAAGTGTCTGCACATCACCGAAGCATGAGAGAG 630  
  
QY      301  tacagcgaacaagtgaagccagcgaactgtgagtcgacgccgcctgcgcgcggag 360  
      |||||||  
Db      631  TACAGGACAAATGAAAGCCACGCTGAGTCGACGCCCGCTCGCGCGCGGAG 690  
  
QY      361  cggctcgaatctggagagcgg-ccggaagccagttaagaagaaggaatgaagaagga 419  
      |||||||  
Db      691  CGGTTGAGTTGGAGACGCTCGGAAAGCCAGTAAAGAACGAGCAGTGCAGCAAGAG 750  
  
QY      420  cctgctgtccgaactacagacatcagcgacagcgagag 458  
      |||||  
Db      751  GGAACGTGCCGCTCGACTTACAGAGGACATCACGACAG 789  
  
RESULT 11  
BI018825/c 493 bp mRNA EST 14-JUN-2001  
LOCUS BI018825 IL3-MT0267-050101-435-H02 MT0267 Homo sapiens CDNA, mRNA sequence.  
ACCESSION BI018825  
VERSION BI018825.1 GI:14425455  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 493)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, R., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
JOURNAL MEDLINE  
20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics
```

```
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?fl=IL3&t2=IL3-MT0267-  
050101-435-H02&t3=2001-01-05&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 13  
High quality sequence stop: 283.  
Location/Qualifiers  
1..493  
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone_lib="MT0267"  
/dev_stage="Adult"  
/note="Organ: marrow; Vector: puc18; Site: 1: SmaI; Site: 2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."  
BASE COUNT      66 a      178 c      90 g      159 t  
ORIGIN  
  
Query Match      47.1%; Score 310.6; DB 11; Length 493;  
Best Local Similarity 91.2%; Pred. No. 9.3e-42;  
Matches 344; Conservative 0; Mismatches 24; Indels 9; Gaps 1;  
  
QY      293  gagaagactacagcagaacgaatgaagcccaactgagtcgcagccgcctgcgcgc 352  
      |||||||  
Db      493  GAGAGATTACAGCAGCAAAAGTGAAGGCCAGATTGGAGTCCAGCCCGCTTGGAGC 434  
  
QY      353  cgcggagcgggttcgagttggagagcggccggaagcagtaaaagaagaatggaag 412  
      |||||||  
Db      433  CGCGGAGCGGTTGAGATTGGAGAGCGCCCGCAAGCCAGTAAAGAGAGCAAAATGGAAG 374  
  
QY      413  aaagggaccctgtctgcacttacagacatcagcgagaggaagaaggaagccagctcg 472  
      |||||||  
Db      373  AAAGGACCTCTGCTCCTTCACAGGCGATCTGACAGCGAGAGAGAACCCAGCTCGG 314  
  
QY      473  ccgaatcctctcagcagaatcaggtctctggaagaagaaggaaggaaggaag 532  
      |||||||  
Db      313  CCGAGTCTCTGCTCATTTGTTATCAGGTCCAGGTTCTGAGGAATTAACAGAGAGAGAG 254  
  
QY      533  aggaagaggaaggaaggaacccagtgaagaatca-----gaggaaggaaggaag 583  
      |||||||  
Db      253  AGGAGAGAGAGAGAGAGGAGCACCAGTGAAGATCAGAGAGAGAGAGAGGAATAGGAG 194  
  
QY      584  aggaagaggaaggaagcgcgcgaacactctggaagggatcagaacagctctgcgaaag 643  
      |||||||  
Db      193  AGGAGAGAGAGAGAGCGCGGACCACTCTGAGAGGATCAGAGCAGTCTCGAAGAAG 134  
  
QY      644  taagtgaagaagaatg 660  
      |||||||  
Db      133  TAACTGAGGAACAATG 117  
  
RESULT 12  
BF607177 830 bp mRNA EST 01-APR-2001  
LOCUS BF607177  
DEFINITION MYL_000119 Mouse 9-day fetus CDNA library ICRP522 Mus musculus  
CDNA clone ICRFP52210930 5', mRNA sequence.  
ACCESSION BF607177  
VERSION BF607177.1 GI:13503669  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus
```

```

/notice="Organ: uterus; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average Insert size 1.75 kb. Life Technologies catalog #:
11538-014"

```

Query Match	50.4%;	Score 332.8;	DB 10;	Length 592;
Best Local Similarity	92.8%;	Pred. No. 2.2e-45;		
Matches 371;	Conservative	0;	Mismatches 2;	Indels 27;
				Gaps 1

RESULT	9
LOCUS	AI580044/c
DEFINITION	AI580044 592 bp mRNA EST 14-DEC-1999
	tct445608 .x1 NCI CGAP Uti Homo sapiens cDNA clone IMAGE:2211806 3'
	similar to TR:061399 Q61399 CELL DIVISION CYCLE 2-LINE 2 ; contains
	element TARI repetitive element ; mRNA sequence.

REFERENCE
1 (bases 1 to 592)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 592)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>

JOURNAL COMMENT
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R Emmert-Buck, M.D., Ph.D.

CDNA library prepared by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

www.dio.liml.gov/bdrp/image/image.htm
Insert length: 1082 Std Error: 0.00
Seq primer: -40UP from Glbco
High quality sequence stop: 405
POLYA-No.

FEATURES	Location/Qualifiers
source	1. .592

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2211806"
/clone_1ib="NCI CGAP D11"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"

```

RESULT 10
BG574409

LOCUS	804 bp	EST	10-APR-2001
DEFINITION	60296352P1 NIH.MGC.87 Homo sapiens cDNA clone IMAGE:4705189 5		
LOCUS	804 bp	EST	10-APR-2001
DEFINITION	60296352P1 NIH.MGC.87 Homo sapiens cDNA clone IMAGE:4705189 5		

mRNA sequence.

ACCESSION	BG574409
VERSION	BG574409.1
	GI:13582062

KEYWORDS	EST.
SOURCE	human.

REFERENCE. 1 (bases 1 to 804)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS NTH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: csagbbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNL10571 row: g column: 14
High quality sequence stop: 686.

FEATURES	Location/Qualifiers
source	1. .804

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:4705189"
```

<hr/>					
RESULT	7				
LOCUS	BF156008	516 bp	mRNA	EST	30-OCT-2000
DEFINITION	RCO-HT0955-270900-033-d08 HT0955 Homo sapiens CDNA, mRNA sequence.				
ACCESSION	BF156008				
VERSION	BF156008.1 GI:11051191				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens Euxaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	Dias Neto,E., Garcia Correa,A., Verjovski-Almeida,S., Britones,M.R., Nagai,M.A., da Silva,W.J.F., Zago,M.A., Bordin,S., Costa,F.F., GOLDMAN,G.H., Carvalho,H.T., Matsukuma,A., Balta,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare, M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags Proc.Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)				
JOURNAL MEDLINE COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel.: +55-11-2704922 Fax: +55-11-2707001 Email: asimpsone@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL: http://www.ludwig.org.br/scripts/gethtml2.pl?tl=bt2-RCO-HT0955-270900-033-d08&t3=2000-09-27&t4=1) Seq primer: puc 18 forward High quality sequence start: 56 High quality sequence stop: 516. Location/Qualifiers 1..516 /organism="Homo sapiens" /db_xref="taxon:9606" /cclone_id="HT0955" /dev_stage="Adult" /note="Organ: head-neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORSTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."				
FEATURES SOURCE					
BASE COUNT	155 a 109 c 175 g 77 t				
ORIGIN					
Query Match	51.4%; Score 339.4; DB 11; Length 516;				
Best Local Similarity	95.8%; Pred. No. 1.9e-46;				
Matches 384; Conservative	0; Mismatches 6; Indels 11; Gaps 3;				
Dy	271	tcttcacatccccaagatgag-agaggactacaagcaaaagtgaaccagcactg	329		
Dy	75	TTCGTCAATCCACCAGCATGAGCAGAGACTTACAGCCAAAGTAAGCACGCACCTG	134		
Dy	330	gagtcgc-agcccgcctcgccgcgcgcgggcgggttcgaattggagaagccgggaac	388		
Dy	135	GAGTCGGTAGTGCGCCTTCGCCGCCGCCGGAGACGGTTGAATTGGAGAACGCCGGAAC	194		
Dy	389	cagtaaaaagaagaaaaatggaagaaagggaacctgctctccgacttacagacatacg	448		
Dy	195	CAGTAAAAGAAGACAATAATGAGAAAAAGGAGCGTGTCCGACTTACAGACATCAACG	254		
Dy	449	aacacgagaggaagacacgcgcgcgcgagctctctgttagcaggaatcagggccagttct	508		

Db	255	ACACCAGAGAGAAGACCACTGCGCCGACTTCTCTGCATCGCGAATTCAGGGTCAGGTCTCTG	314
OY	509	aggagaagaggggagggaggaagaggaaggaaggaaggaaggaaggaagcaatcgtgaagaaatca-	567
Db	315	AGGAAGAAGAGGAGGAGGAGAAAAAGAGAGAGAGAGAAAGGAGGAGAACCAAGTCAGAAATCAG	374
OY	568	-----gagagaggaagaggaaggaaggaaggaagccgcacccaactctgaagag	619
Db	375	AGGAGGAGGAGGAGGAGGAAGGAAGGAGAGAGAGAGACAACCGCACGCACTCTGAGSAGG	434
OY	620	cattcagagcagtcctgccgaagaagtgaatgtaaggaagaatg	660
Db	435	CATCAGAGCAGTCTGCCGAAGAAGTAAGTGAGCAGCAAGATG	475
RESULT	8		
LOCUS	BC009375	906 bp mRNA	HMC 09-JUL-2001
DEFINITION	Homo sapiens, Similar to cell division cycle 2-like 1 (PITSLRE proteins), clone IMAGE:4121554, mRNA.		
ACCESSION	BC009375		
VERSION	BC009375.1	GI:14627288	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 906)		
JOURNAL	Strausberg, R. Direct Submission Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: gcgabs@email.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamins, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Masello, C., Mastrian, S.D., McCloskey, D.C., McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgou, C., Vogt, J.T., Walker, M.A., Zhang, L.-H. and Green, E.D.		
FEATURES			
source	Location/Qualifiers 1..906 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4121554" /rname_type="muscle, Thadomyosarcoma" /clone_id="NTL_MGC_17" /lab_host="DH10B-R" /note="Vector: pOTB7"		
BASE COUNT	317 a 170 c 307 g 112 t		
ORIGIN			
Query Match	51.0%, Score 336.6, DB 12, Length 906;		

Query Match 71.9%; Score 474.8; DB 10; Length 1042;
 Best Local Similarity 89.0%; Pred. No. 1.4e-68;
 Matches 596; Conservative 0; Mismatches 32; Indels 42; Gaps 6;

```

QY 5 aaagtcggaagacactctgagaagaacagataaagctcgccgggaatgggaagacaga 64
    |||||||
Db 1 AACGTGCGAAMCGCATCAGAAAGAACAGATTAACCTCGCGGAAATGGAAAGACAGA 60
    |||||||
QY 65 agagaagggaagaatgggaagg-gagcattccagagaagaaggggatgatgtgctgtgc 123
    |||||||
Db 61 AGGAAGAGGAGATGGCAAGGTGAGCATTCACAGGAGGAA----- 99
    |||||||
QY 124 ctcttcaggaagccgcttggagcagttagaaagaacgaggagcgaggacgcaagatcgcg 183
    |||||||
Db 100 -----AGGAGCCGCTTGGAGCAGTTAGAAAGAACCGGGAGCGGAGCAAGATCGCG 153
    |||||||
QY 184 gagcagcagaagaagcagcgggagcagaagaagcgcgagcgcgcgaggaagcgcgcg 243
    |||||||
Db 154 GACACACAGAGAGAGAGCGGGGAGCAAGAGAGCCGAGCGGCGGAGAGAGCGCGCGC 213
    |||||||
QY 244 aaggaagcggagagcccgacggaggaagtgtctgcacatcacggaacgatatgagagactac 303
    |||||||
Db 214 AAGGAGCGGAGAGCCCGCAGGGAAGTGTCTGCACATCACGAACGATGAGAGGACTAC 273
    |||||||
QY 304 agcgacaagaatgaaagccacgtgagtcgcaagcccgctcgagcgcgcgcgagagcg 363
    |||||||
Db 274 AGCGACAAAGTGAAGCCAGCCACTGTGAGTCCGACGCCGCTCGCGCGCGGAGCGG 333
    |||||||
QY 364 ttggaattgggaagcggcggaagccaagtaa-----aagaagaagaatggaaagaagga 419
    |||||||
Db 334 TTGAGATTGGGAGACGGCCGGAAGCCAGTAAACGAAGAAACTGTGGCAAGCAAGAGGA 393
    |||||||
QY 420 cctgctgcctcgaactcaagagacatcagcgacagcgaggaagaagcaagctcgccgagtc 479
    |||||||
Db 394 CCGGCTGTCCGACTTACAGAGACATCAGCGACAGAGAGAAAGACACGCTCGCGCGAGTC 453
    |||||||
QY 480 ctctgcaagcaaatcagagctcaggtctctgagaagaagaaggaaggaaggaag--agag 537
    |||||||
Db 454 CTGCTTACGCGCAATCAGGCTCAGGTTCTGTGGAAGAAAGAGAGAGAGCAAGCAAGAG 513
    |||||||
QY 538 gaaggaagaagaagcagccagttgaagaatcagagaagaagaagaagaagaaggaagga 596
    |||||||
Db 514 ACGGAGAGCGGAGACACCACTGAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
    |||||||
QY 597 -----gaccgagcagaactctgagagagcatcagagcagtcgcggaagaagtaagtg 650
    |||||||
Db 574 GAGGAGGAGACGAGAGAACTCTGAGGAGGCAATCAGAGCAAGCTTG--CGAAGAACTAAGTGA 632
    |||||||
QY 651 ggaagaagaatg 660
    |||||||
Db 633 GGAAGAAATG 642
    |||||||

```

RESULT 5
 LOCUS BG822824 662 bp mRNA
 DEFINITION 602727816f1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4867443 5',
 mRNA sequence.
 ACCESSION BG822824
 VERSION BG822824.1 GI:14170411
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 1 (bases 1 to 662)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaps-f@mail.nih.gov

Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: NIH Intramural Sequencing Center
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: L10M1735 row: p column: 04
 High quality sequence stop: 601.
 Location/Qualifiers
 1. 662
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4867443"
 /clone_1db="NIH_MGC_15"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: colon; Vector: pORF7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAAGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 202 a 139 c 255 g 66 t
 ORIGIN

Query Match 63.6%; Score 420; DB 11; Length 662;
 Best Local Similarity 99.5%; Pred. No. 1.3e-55;
 Matches 442; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```

QY 188 agcagaagaagcaagcgggagcagaagaagcgcgagcgcgaggagagcgagcgcaag 247
    |||||||
Db 2 AGCAAGAAGAGAGAGCGGAGAGCAAGAGAGCGGAGCGGCGAGAGCGGCGCAAGG 61
    |||||||
QY 248 agcgaggagcccgaggggaagtgtctgcacatcacggaagatgagagagactaacgcg 307
    |||||||
Db 62 ACCGGGAGGCGCGGAGAGATGTCTGCACATCACGAAAGATGAGAGAGACTACAGCG 121
    |||||||
QY 308 acaaatgaagaagccagcagctgagctgcagcccgccctcgccgagcgagcggttcg 367
    |||||||
Db 122 ACAAAAGTAAACCCAGCCACTGTGAGTCCGACCCGCTCGGCGCGGAGAGCGGTTCCG 181
    |||||||
QY 368 agttggagagcggccggagagcagtaaaagaagaagaatgaaagaagagagcctgtct 427
    |||||||
Db 182 AGTTGGAGAGCGCGCGGAGAGCACTTAAGAAAGAAATGGAAGAAAGAGCACTGCTGT 241
    |||||||
QY 428 ccgacttaagagacatcagcgacagcgagaggaagaccagctcgcgagctcctctcag 487
    |||||||
Db 242 CCGACTTACAGGACATCAGCGACGAGAGAGAGACCACTCGGCGCGAGCTCTGTCAG 301
    |||||||
QY 488 cagaatcaggtcaggtctctgagaagaagaggaggaaggaagaaggaaggaaggaag 547
    |||||||
Db 302 CAGAAATCAGGCTCAGGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
    |||||||
QY 548 ggaagcacc-agtgaagaatca-gaggaaggaaggaaggaaggaaggaaggaagcagcg 605
    |||||||
Db 362 GGAGCAGCGAGTGAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
    |||||||
QY 606 caactctgagagagcatcagaagca 629
    |||||||
Db 422 CAACTCTGAGGAGGATCAGAGCA 445
    |||||||

```

RESULT 6
 LOCUS AK008283 1000 bp mRNA
 DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length
 enriched library, clone:2010016J05, full insert sequence.
 ACCESSION AK008283
 VERSION AK008283.1 GI:12842370

QY	486	aggatcagagcctcaggtctctaggaaga-----gaagagagaggaagaagaagg	541
Db	454	AGCGGATCATGCTCAGGCTTCTGAGAAAGACAGAGAGGCGGAAACACAGAGAGAC	513
QY	542	aggaaggagagaccagtgaagaatca-----gaggaagaagaagaagaaggag	592
Db	514	GGACAGAGGAGACCAAGTACGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG	573
QY	593	aggagagccgcagcaactctgagagagca-taagagcagctgcgcgaagaagtatgag	651
Db	574	AGGAGACCGGACGACACTCTGAGAGAGGACATTCAGAGACACTGCCGAAGAAGTAAGTAG	633
QY	652	gaagaatg 660	
Db	634	GAGCAATG 642	
RESULT	3		
LOCUS	Bg107702	1340 bp	EST
DEFINITION	6022277878F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4365474 5',		
ACCESSION	Bg107702		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1340)		
JOURNAL	NIH-MGC http://imgc.nhl.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC Clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.llnl.gov		
	Plate: LHAM0014 row: 1 column: 19		
	High quality sequence stop: 667.		
FEATURES	Location/Qualifiers		
SOURCE	1..1340		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4365474"		
	/clone_1ib="NIH_MGC_86"		
	/tissue_type="osteosarcoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: bone; Vector: pCMV-SPORT6, site.1: NotI;		
	site.2: SalI; Cloned unidirectionally; oligo-dT primed.		
	Average insert size 1.533 kb. Library enriched for		
	full-length clones and constructed by Life Technologies.		
	Note: this is a NIH-MGC Library."		
BASE COUNT	467 a 272 c 450 g 151 t		
ORIGIN			
Query Match	72.2%; Score 476.4; DB 11; Length 1340;		
Best Local Similarity	98.0%; Pred. NO. 7,7e-69;		
Matches 496; Conservative	0; Mismatches 1; Indels 9; Gaps 1		
QY	164	agcgagcgcaagatgaggagcagcagaagagagcagcgaggagcgagc	223
Db	1	AGCGGAGCGCAAGATGCGGAGCAGCAGAAAGACACGCGAGCAGAGCGCGAGC	60
QY	224	ggcggggagagagcgaggcaggaagcgaggagccgcgaaggagtgctcgaatacc	283
Db	61	GCGGCGGAGAGAGCGGCGGAAAGACCGGAGAGCCCGCAGAGGAAGTGTCTTCACATAC	120

QY	284	gaaagatgagagaggaactacacgacgaacaaatgtaaaagccagccactgtgagtcgagccgc	343
Db	121	GAACGATGAGAGGAGCATACACCGCAAAAGTGAAGCCAGCCACTGGATGCGCAGCCCGC	180
QY	344	cctggccgcgcgcgggaagcggttcctgagttctgggagacgcgcggaagcccaataaagaaga	403
Db	181	CTCGGCGCGCGGGAGCGGCTTCGAGTTGGGAGACGCGCCGGAAGCCAGTAAAGAAAGAGA	240
QY	404	aaatggaagaaaggagaccctgctcgttcgaacttaccaggaatcaagcagacgagaggaaga	463
Db	241	AAATGGAAGAAAGGAGACTGCTGCTCGACTTACAGGACATCAGGAGACAGCGAGAGAAAGA	300
QY	464	ccagctcgccgagtcctcctcgttcacagaatacagctcaggttctgtgagaagaagaagg	523
Db	301	CCAGCTCGCGCCGAGTCCCTGTCAGCGGATCAGGCTCAGGTTCTGAGGAAGAAAGAGGAGG	360
QY	524	aaggagaagagaagagaagagaagaagagacacacagtgaaagatca-----gaggaag	574
Db	361	AGGAGGAAGAGAGAGAGAGAGAGAGGAGACACCAGTGAAGATCAGAGAGAGAGAGAGAGG	420
QY	575	aagagaagagaagagaagagaagagaccggaacgaactcgtgagaagcatcagaagcagtcg	634
Db	421	AAGAGGAGAGAGAGAGAGAGAGACCGGCGACACACTCTGAGAGAGGCAATCAGAGCAGTCTG	480
QY	635	ccgaagaagtaagtgagaagaagatg 660	
Db	481	CCGAAGAAGTAAGTGAAGAGAAATG 506	
RESULT	4		
LOCUS	BE742943	1042 bp mRNA	EST 15-SEP-2000
DEFINITION	601574925f1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3835686 5', mRNA sequence.		
ACCESSION	BE742943		
VERSION	BE742943.1	GI:10157033	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1042)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: DCTD/DTF		
	CNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at: image.lnl.gov		
	Plate: L10M518 row: D column: 07		
	High quality sequence stop: 667.		
FEATURES	Location/Qualifiers		
source	1. 1042		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3835686"		
	/clone_id="NIH_MGC_9"		
	/tissue_type="adenocarcinoma cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: ovary; Vector: pOM57; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."		
BASE COUNT	371 a 159 c 390 g 122 t		
ORIGIN			

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:14:23 : Search time 3881.49 Seconds
(without alignments)
1827.188 Million cell updates/sec

Title: US-09-915-060-6

Perfect score: 660
Sequence: 1 caggaacgcgaagaacgaca.....aagtaagtcgaggaagaatg 660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estln:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_estl:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_hov:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vtl:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	568.2	86.1	1063	10	BE740559 601595656
2	518.2	78.5	723	10	BE742222 601575629
3	476.4	72.2	1340	11	BG107702 602277878
4	474.8	71.9	1042	10	BE742943 601574925
5	420	63.6	662	11	BG822824 602727816
6	362.8	55.0	1000	12	AK008283 Mus muscu
7	339.4	51.4	516	11	BF156008 RC0-HT095
8	336.6	50.4	906	12	BC009375 Homo sapi
9	332.8	50.4	592	10	AT580044 tq45g08.x
10	314.2	47.6	804	11	BG574409 602596352
11	310.6	47.1	493	11	BI018825 IL3-MT026
12	286.4	43.4	830	11	BF607177 MY1_00011

13	277.6	42.1	812	11	BT149024
14	271.6	41.2	326	11	BG943585
15	266	40.3	758	11	BG969359
16	265.8	40.3	826	11	BG034678
17	256.6	38.9	372	10	AL034747
18	244.4	37.0	740	10	AL043340
19	212.6	32.2	367	10	AA681955
20	208.2	31.5	345	11	BT261348
21	205.6	31.2	674	10	BE255507
22	170.8	25.9	649	10	AM210026
23	167.2	25.3	283	10	AM804794
24	166.8	25.3	251	10	AV270235
25	166.6	25.2	476	10	BE680738
26	164.4	24.9	900	11	BG298479
27	155.6	23.6	1083	11	BG770237
28	153	23.2	537	11	BF743770
29	149.2	22.6	253	10	AV327502
30	142.2	21.5	292	10	AV368285
31	138	20.9	441	10	AW451883
32	135.8	20.6	233	10	BR290670
33	135.6	20.5	182	10	AA413247
34	134.2	20.3	456	11	BF998861
35	127.6	19.3	725	11	BF308197
36	126.4	19.2	216	10	AV105652
37	124	18.8	514	10	AT989489
38	120.8	18.3	462	11	BG486162
39	119.4	18.1	732	10	AT006623
40	110.2	16.7	160	10	AM193827
41	107.6	16.3	1005	10	AL523189
42	106	16.1	673	11	BF983533
43	104.4	15.8	469	11	BF359805
44	104	15.8	339	11	H33616
45	103.8	15.7	321	10	AA309711

ALIGNMENTS

RESULT 1
LOCUS BE740559 1063 bp mRNA
DEFINITION 601595656F1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3949693 5',
mRNA sequence.
ACCESSION BE740559
VERSION BE740559.1 GI:10154551
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1063)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DMP

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMU at: image.llnl.gov
Plate: L10M814 row: p column: 14
High quality sequence stop: 741.
Location/Qualifiers

FEATURES

source
1..1063
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949693"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:

Qy 628 cagtcctgcccgaagaagtaagtgaagaaga 656
| | | | | | | | | | | | | | | | | |
Db 106 GGGCAGGAGGAGGAGGAGGAGGAGGAGGA 78

Search completed: December 4, 2001, 18:58:37
Job time: 9263 sec

```
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSK/SKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3489 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3489
US-08-728-323A-1
```

```
Query Match 11.4%; Score 75.4; DB 2; Length 3489;
Best Local Similarity 46.6%; Pred. No. 2.2e-07;
Matches 241; Conservative 0; Mismatches 276; Indels 0; Gaps 0;
```

```
OY 141 ggaagcagttgaagaagcggagcgaggcaagatgcggagcagcagaagaagca 200
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2172 GCAGCAGGATGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 2231
OY 201 gcggagcagaagaagcgcgagcgcgagcgagcgagcgagcgagcgagcgagcgagcg 260
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2232 GGATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 2291
OY 261 caaggagagtgctgcacatccgaacgatgagaagaactacagcgacaagtgaagc 320
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2292 GCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 2351
OY 321 cagccactgagtcgagcccgctcgcccgccgagcgagcgagcgagcgagcgagcgagcg 380
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2352 GGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTT 2411
OY 381 ccggaagcagtaaaagaagaagaatggaagaagaaggaactgctgcgacttaacaga 440
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2412 AGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 2471
OY 441 catcagcagcagagagagagacccagctcgcccgagtcctgcagcagaatcaggtc 500
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2472 GTTAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGGT 2531
OY 501 aggtctggaagaagaagagagagagagagagagagagagagagagagagagagagagag 560
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2532 GGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTA 2591
OY 561 agaatcagaaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 620
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2592 GGAGGTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 2651
OY 621 atcagagcagtcgcccgaagaagtaagtgaagaaga 657
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2652 GGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTA 2688
```

```
RESULT 13
US-08-770-379-20/c
Sequence 20, Application US/08770379
Patent No. 5849564
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
```

```
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/770,379
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 52342
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ. ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 32207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-770-379-20
```

```
Query Match 11.4%; Score 75.4; DB 2; Length 32207;
Best Local Similarity 46.6%; Pred. No. 2.7e-07;
Matches 241; Conservative 0; Mismatches 276; Indels 0; Gaps 0;
```

```
OY 141 ggaagcagttgaagaagaagcggagcgaggcaagatgcggagcagcagaagaagca 200
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 19825 GCAGCAGGATGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 19766
OY 201 gcggagcagaagaagcgcgagcgcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 260
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 19765 GGATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 19706
OY 261 caaggagagtgctgcacatccgaacgatgagaagaactacagcgacaagtgaagc 320
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 19705 GCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTA 19646
OY 321 cagccactgagtcgagcccgctcgcccgccgagcgagcgagcgagcgagcgagcgagcgagcg 380
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 19645 GGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTT 19586
OY 381 ccggaagcagtaaaagaagaagaatggaagaagaaggaactgctgcgacttaacaga 440
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 19585 AGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTA 19526
OY 441 catcagcagcagcagagagagagagagagagagagagagagagagagagagagagagagagag 500
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 19525 GTTAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGGT 19466
OY 501 aggtctggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 560
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 19465 GGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTA 19406
OY 561 agaatcagaaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 620
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 19405 GGAGGTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGT 19346
OY 621 atcagagcagtcgcccgaagaagtaagtgaagaaga 657
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 19345 GGAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTA 19309
```

QY 261 cagggaagtgctgcacataccgaacgatgagagagactacaagcgacaagtgaagc 320
Db 2603 GGACGACGAGGGGCGACGAGGCGAGAGGGGCGACGAGGGGCGACGAGGCGACGAGGA 2662
QY 321 cagccactgagtgctcagcccgccctcgcccgcccgagagcggttcagttgagagacg 380
Db 2663 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2722
QY 381 ccggaagccagtaaaagaagaataatggaagaagaaggaacctgtctgcagacttaacga 440
Db 2723 GGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 2782
QY 441 catcagagcagcgagagagagagacacgtcgccgagtcctcgtccagcagaatcaggtc 500
Db 2783 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 2842
QY 501 aggttcgagaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 560
Db 2843 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 2902
QY 561 agaatacgaagagagagagagagagagagagagagagacggcagcaactctgagagac 620
Db 2903 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2962
QY 621 atcagagcagctgtcggaagaagaatgaagtgaag 652
Db 2963 AGTGGAGGCCGGGGTCTGAGGAGTGTGAGG 2994

RESULT 11

PCT-US93-04648-15
Sequence 15 Application PC/TUS9304648

GENERAL INFORMATION:

APPLICANT: Genentech, Inc., Godowski, Paul J., Lokker, Natalie A., Mark, Melanie H.
TITLE OF INVENTION: HEPATOCTYTE GROWTH FACTOR VARIANTS

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: palin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/04648

FILING DATE: 19930517

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/884811

FILING DATE: 18-MAY-92

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/885971

FILING DATE: 18-MAY-92

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Ginger R.

REGISTRATION NUMBER: 33,055

REFERENCE/DOCKET NUMBER: 755,779P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-3216

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 10596 bases

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

PCT-US93-04648-15

Query Match 11.9%; Score 78.4; DB 5; Length 10596;
Best Local Similarity 47.1%; Pred. No. 6e-08;
Matches 241; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 141 gggagtgtagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 200
Db 2483 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 2542
QY 201 gcggaagcagaagaagcgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 260
Db 2543 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 2602
QY 261 cagggaagtgctgcacataccgaacgatgagagagagactacaagcgacaagtgaagc 320
Db 2603 GGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 2662
QY 321 cagccactgagtgctcagcccgccctcgcccgcccgagagcggttcagttgagagacg 380
Db 2663 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 2722
QY 381 ccggaagccagtaaaagaagaataatggaagaagaagaagaacctgtctgcagacttaacga 440
Db 2723 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2782
QY 441 catcagagcagcgagagagagagacacgtcgccgagtcctcgtccagcagaatcaggtc 500
Db 2783 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 2842
QY 501 aggttcgagaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 560
Db 2843 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGA 2902
QY 561 agaatacgaagagagagagagagagagagagagagagacggcagcaactctgagagac 620
Db 2903 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2962
QY 621 atcagagcagctgtcggaagaagaatgaagtgaag 652
Db 2963 AGTGGAGGCCGGGGTCTGAGGAGTGTGAGG 2994

RESULT 12

US-08-728-323A-1
Sequence 1 Application US/08728323A

Patent No. 5948676

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Russo, James J.

APPLICANT: Edelman, Isidore S.

APPLICANT: Moore, Patrick S.

TITLE OF INVENTION: Immediate Early Protein From Kaposi's

TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/728,323A

FILING DATE:

CLASSIFICATION: 435


```

; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELE: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-885-971-15

```

```

Query Match      11.9%; Score 78.4; DB 1; Length 10596;
Best Local Similarity 47.1%; Pred. No. 6e-08;
Matches 241; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

```

```

QY 141 ggaagcagttagaagaagcggagcggagcgaagaatgcyggaagcagcagaagagca 200
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2483 GGGGCGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2542
QY 201 gcggagcagaagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 260
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2543 GGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2602
QY 261 cagggaagtgctgcacatcccaacgaatgagaagagactacgcgcacaatggaagc 320
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2603 GGAGCGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2662
QY 321 cagccacgcgagtcgcagccgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 380
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2663 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2722
QY 381 ccggaagcagtaaaagaagaagaatggaagaagagcctgtctgcgaacttacaga 440
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2723 GGAGCGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2782
QY 441 catcagcagcagcagaggaagaccagctcgccgagtcctcgtcagcagaatcagctc 500
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2783 GGGGCGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2842
QY 501 aggttcctgagaagaagaagaggaagaagaagaagaagaagaagaagaagaagaaga 560
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2843 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2902
QY 561 agaatcagagaggaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 620
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2903 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2962
QY 621 atcagagcagtcgtccgagaagaatgagc 652
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2963 AGTGGAGCCCGGGGTGAGGAGGTGAGGAGG 2994

```

```

RESULT 8
US-08-087-783A-15
; Sequence 15, Application US/08087783A
; Patent No. 5547836
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.

```

```

; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,783A
; FILING DATE: 13-Jul-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0755779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELE: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: linear
US-08-087-783A-15

```

```

Query Match      11.9%; Score 78.4; DB 1; Length 10596;
Best Local Similarity 47.1%; Pred. No. 6e-08;
Matches 241; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

```

```

QY 141 ggaagcagttagaagaagcggagcggagcgaagaatgcyggaagcagcagaagagca 200
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2483 GGGGCGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2542
QY 201 gcggagcagaagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 260
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2543 GGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2602
QY 261 cagggaagtgctgcacatcccaacgaatgagaagagactacgcgcacaatggaagc 320
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2603 GGAGCGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2662
QY 321 cagccacgcgagtcgcagccgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 380
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2663 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2722
QY 381 ccggaagcagtaaaagaagaagaatggaagaagagcctgtctgcgaacttacaga 440
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2723 GGAGCGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2782
QY 441 catcagcagcagcagaggaagaccagctcgccgagtcctcgtcagcagaatcagctc 500
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2783 GGGGCGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2842
QY 501 aggttcctgagaagaagaagaggaagaagaagaagaagaagaagaagaagaagaaga 560
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2843 GGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2902
QY 561 agaatcagagaggaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 620
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2903 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2962

```


Matches	241: Conservative	0: Mismatches	271: Indels	0: Gaps
Qy	141	ggagcagcttagaagaagacggyagacggyagacgaagtgcggyagacgaagaagagca	200	
Dy	141			
Dy	949	ggggcagagacagacagagagggcagagggcagagcagacagacagagagggcagagggcgca	1008	
Qy	201	ggcgagagcagaagagacgcyagacggyagcgygagagcgycgcgaaggagcgygagagccg	260	
Dy	1009	ggagggcgacagacagacagagagggcgagagcagacagagggcgacagagggcgacagagggcgca	1068	
Qy	261	cagggaagtgctctgcacatccacgaacgcatgagaaagactacagcgacaaagtgaagc	320	
Dy	1069	ggagcagagagggcgacagacagacagagggcgacagggcgacagcagcagcagcagcagcagcagca	1128	
Qy	321	cgacgcctggaatgcgagcccgccctcgccgcccggggagagaggtctggaatgttgagaacg	380	
Dy	1129	ggggcagagacacagagagggcgacagacagacagagggcgagagagagagagggcgacagggggcgca	1188	
Qy	381	ccggaagccacgtataaagaagaanaatlgaaanaaaggaacctgtctgcgactaccagga	440	
Dy	1189	ggagcagagagggcgacagacagacagagggcgacagagggcgacagagggcgacagagggcgacagagga	1248	
Qy	441	catcagcgaagcggagaggaagacccaagctcgcgagatccctgcagcagatcaggtc	500	
Dy	1249	ggggcagagagggcgacagacagacagagggcgacagggcgacagagggcgacagagggcgacagagggcgca	1308	
Qy	501	aggtctcgaagaagaagagaggaaggaagaagaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga	560	
Dy	1309	ggggcagagacagacagagagggcgacagagggcgacagagggcgacagagggcgacagagggcgacagagga	1368	
Qy	561	agaatcagaaggaagaagaagaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga	620	
Dy	1369	ggagagagggcgacagacagacagagggcgacagagggcgacagagggcgacagagggcgacagagggcgacagagga	1428	
Qy	621	atcagagcagctctccgaagaagtgaagtgaag	652	
Dy	1429	agtgagagggcgggggtcgagagaggtactgagag	1460	
RESULT 6				
US-07-884-811-15				
Sequence 15, Application US/07884811				
Patient No. 5316921				
GENERAL INFORMATION:				
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.				
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCTYTE GROWTH FACTOR VARIANTS				
NUMBER OF SEQUENCES: 21				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: Genentech, Inc.				
STREET: 460 Point San Bruno Blvd				
CITY: South San Francisco				
STATE: California				
COUNTRY: USA				
ZIP: 94080				
COMPUTER READABLE FORM:				
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk				
COMPUTER: IBM PC compatible				
OPERATING SYSTEM: PC-DOS/MS-DOS				
SOFTWARE: palin (Genentech)				
CURRENT APPLICATION DATA:				
APPLICATION NUMBER: US/07/884,811				
FILING DATE: 19920518				
CLASSIFICATION: 530				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER:				
FILING DATE:				
ATTORNEY/AGENT INFORMATION:				
NAME: Dreger, Ginger R.				
REGISTRATION NUMBER: 33,055				
REFERENCE/DOCKET NUMBER: 755.1				
TELECOMMUNICATION INFORMATION:				
TELEPHONE: 415/225-3216				
TELEFAX: 415/952-9881				

```

; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-884-811-15

Query Match      11.9%; Score 78.4; DB 1; Length 10596;
Best Local Similarity 47.1%; Pred. No. 6e-08;
Matches 241; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 141 ggaagcaagttaagaagaagcggagcggagcgaagaatgcgggagcagcaagaagagca 200
    |||||
Db 2483 GGGGCGAGGACGACGAGAGGAGGGGCGAGAGGGGCGAGAGAGGGGCGAGAGAGGGGCA 2542
QY 201 gcgaggagcagaagaagcgcgagcggcgcgagagagcgcgcaagagcgagagcccg 260
    |||||
Db 2543 GGAGGGGGCGAGAGCAAGAGAGGGGGCGAGGAGCGAGAGGGGCGAGAGGGCGAGAGGGGCA 2602
QY 261 cagggagatgtctgcacatccaccgaagatgagagagagatcagcgacaaagtgaagc 320
    |||||
Db 2603 GGAGCAGAGGAGGGGCGAGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCAGCA 2652
QY 321 cagccacgtgagctgcagcccgcccgcccgcccgcgagcgagcttgagatgtgagagcg 380
    |||||
Db 2663 GGGGCGACGAGCAGAGAGGGCGAGAGGCGCAGAGAGGGGCGAGAGAGGGGCGAGAGGGGCA 2722
QY 381 ccggaagccagtaaagaagaagaatgagaaagaggaacctctgtctccgacttacagga 440
    |||||
Db 2723 GGAGCAGAGGCGGGCGAGAGGGCGAGAGCAGAGGCGCAGAGAGGGGCGAGAGCAGAGCAGCA 2782
QY 441 catcagcgacagcagcgagaggaagaaccaagctcgggccgagctctcgtaagcaaatcaagctc 500
    |||||
Db 2783 GGGGCGAGGAGGGGCGAGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGAGGGGCGAGAGCA 2842
QY 501 aggttcctgaggaagaagaagagagagagagaaagaagaagaagaagaagagcaccagtga 560
    |||||
Db 2843 GGGGCGAGGAGCGAGAGGGCGCAGAGGGGCGAGAGCGAGAGGGGCGAGAGGGGCGAGAGGCA 2902
QY 561 aqaatcagaagagagagagagagagagagagagagagagagacggcagagcaactctgagagagc 620
    |||||
Db 2903 GGAGCAGGGGGCGAGAGCAGAGGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGCG 2962
QY 621 atccagagcagctctgccgaagaaglaagtagc 652
    |||||
Db 2963 AGTGAAGCGCGGGGCTCGAGAGAGGTAGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2994

RESULT
US-07-885-971-15
; Sequence 15, Application US/07885971
; Patent No. 5328837
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885,971

```

[illegible]

Query Match	11.9%;	Score 78.4;	DB 2;	Length 5452;
Best Local Similarity	47.1%;	Pred. No. 5.6e-08;		
Matches 241; Conservative	0;	Mismatches 271;	Indels 0;	Gaps 0;

[illegible]

OY	261	cagggaagttctctgcaatattaccggaacgatggaagagagctatcacgcacaaagtgtaaagc	320
Db	1782	GGACACAGGAGGGCGCAGGAGCAGGAGGAGGGCGCGAGGGGCGAGAGGGCGACAGACAGGA	17233
OY	321	cagcacactggaatctgcagcccgccctcgcccgcccgcggaagcggtctcgagtttbgaggaacg	380
Db	1722	GGGGCAGGAGCAGGAGGCGCAGGAGCAGGAGGGGCGAGAGCGAGAGGGGCGAGAGGGCA	16637
OY	381	ccggaagccagtaaaagaagagaanaatggaagaagggagacctgtctccgattacagga	440
Db	1662	GGACACAGGAGGCGCAGGAGGCGCAGGAGCAGGAGGGGCGAGAGGGGCGACAGACAGAGGA	16030
OY	441	catcaaggaagagcgagagggagagaccgagctctgcccaggtctctctgtaagcagaatcaagctc	500
Db	1602	GGGGCAGGAGGGCGCAGGAGCAGGAGGAGGGGCGAGGGGCGAGCAGAGGAGGGGCGACGA	15433
OY	501	aggtctctgagaaagaagaaggaaggaagaagaaggaaggaaggaaggaaggaaggaacacgtga	560
Db	1542	GGGGCAGGAGCAGGAGGCGCAGGAGGGGCGAGGAGCGAGAGGGGCGAGAGGGGCGAGAGCA	14833
OY	561	agaaatcagaggaaggaaggaaggaaggaaggaaggaaggaacccgacagcaactctgaagaaagc	620
Db	1482	GGAGGAGGGGCGAGGAGCAGGAGGCGCAGGAGCAGGAGGTGGAGGCCGGGGGTGAGAGAGGC	14233
OY	621	atcagagcagttctgcgaagaagaatgaagtgg	652
Db	1422	AGTGAAGGCCGGGGTGCAGGAGGTAGTGAAGC	1391

RESULT 5
US-08-910-647-1

; Patent No. 6251433

GENERAL INFORMATION:

APPLICANT: Zuckermann et al

[illegible]

NUMBER OF SEQUENCES.	TITLE OF INVENTION:	PC
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

; CITY: Emeryville

STATE: California

COUNTRY: U.S.A.
STD: 04609-2016

ZIP: 94608-2916

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0,

```

; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/910,647

CLASSIFICATION: 514

CLASSIFICATION: 314
ATTORNEY/AGENT INFORMATION

NAME: Fujita, Sharon M.

REGISTRATION NUMBER: 38,459

REFERENCE/DOCKET NUMBER: 1218.002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 923-2706

TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 1.

```

; INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 9600 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE 1

Query Match	11.98;	Score 78.4;	DB 4;	Length 9600;
Best Local Similarity	47.18;	Pred. No. 5.9e-08;		

```
Query Match          15.1%; Score 99.4; DB 1; Length 7218;
Best Local Similarity 5.4%; Pred. No. 3e-12;
Matches 22; Conservative 259; Mismatches 130; Indels 0; Gaps 0;

QY 2 acgaagtcggaacacgacgcgaagaagaacagataagctcgccggaatggaagac 61
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1469 ATGCAAGTAGTTAAGAGATAGAGATTGGTACRRRRRRRRRRRRRRRRRRRRR 1410

QY 62 agaagaagaagaatgcaaggaagcatccaggaagaagaagaaggaatgagctgt 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1409 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1350

QY 122 gcccttcgaagcgcttggaagcagtaagaagaacggaagcggaagcgaatgc 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1349 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1290

QY 182 ggaagcagcagaagaagcagcggaagcagaagaagcgcgcgcgcgcgcgcgc 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1289 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1230

QY 242 gcaagagcggaagcgccgcgaagtgctgcacatccgaacgatgagaagact 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1229 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1170

QY 302 acagcgacaagtgaaagccagccactgagctgcagccgcgcgcgcgcgcgcgc 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1169 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1110

QY 362 ggttcgagtcggaagcgccgaagcagcagtaagaagaagaatggaag 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1109 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1059

RESULT 2
US-09-130-114-2/c
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Dama, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/ID9030US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

Query Match          12.8%; Score 84.8; DB 2; Length 1931;
Best Local Similarity 46.6%; Pred. No. 2.5e-09;
Matches 272; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

QY 72 ggaatgcaagggagcatccaggaagaagaagggaatgagctgctcttcag 131
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 953 GGACGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 894

QY 132 ggaacgcttgagcagcttgaaagaagcggaagcggaagcgaagaatgcagca 191
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 893 GGACGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 834

QY 192 gaagagcagcggaagcagaagagcgcgagcggaagcggaagcggaagcg 251
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 833 CGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774

QY 252 ggaagcccgaggaagctctgcacatccgaacgatgagaagagactacagcgaca 311
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 773 CGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
QY 312 agtgaagccagccactgagtcgagccgcctcgccgcgcgcgcgcgcgcgcgc 371
Db 713 GGACGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
QY 372 ggaagacgagcggaagccagtaagaagaagaatggaagaagaagccgctgcga 431
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 653 CGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594
QY 432 cttaacagacatcagcagaagcagaagaagacacagctcgccgagctcctgcagcaga 491
Db 593 CGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534
QY 492 atcagagctaggttctggaagaagaagagagagagagagagagagagagag 551
Db 533 GGACGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474
QY 552 caacagtgagaatcagaaggaaggaagaagaaggaaggaagcgcgcgcgcgcgc 611
Db 473 GGACGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 414
QY 612 tgaagagcatcagaagcagctgcggaagaagtaagtgaggaag 655
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 413 GGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370

RESULT 3
US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hwang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-050-863-2

Query Match          11.9%; Score 78.4; DB 3; Length 2580;
Best Local Similarity 47.1%; Pred. No. 5.1e-08;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:57:53 ; Search time 152.61 seconds
(without alignments)
979.461 Million cell updates/sec

Title: US-09-915-060-6

Perfect score: 660

Sequence: 1 cagcaagctcggaagacaca.....aagtaagtgaagaaatg 660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 segs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
C 1	99.4	15.1	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	84.8	12.8	1931	2 US-09-130-114-2	Sequence 2, Appl
C 3	78.4	11.9	2580	3 US-09-050-863-2	Sequence 2, Appl
C 4	78.4	11.9	5452	2 US-09-130-114-1	Sequence 1, Appl
C 5	78.4	11.9	9600	4 US-08-910-647-1	Sequence 1, Appl
C 6	78.4	11.9	10596	1 US-07-884-811-15	Sequence 15, Appl
C 7	78.4	11.9	10596	1 US-07-885-971-15	Sequence 15, Appl
C 8	78.4	11.9	10596	1 US-08-087-783A-15	Sequence 15, Appl
C 9	78.4	11.9	10596	1 US-08-194-088B-15	Sequence 15, Appl
C 10	78.4	11.9	10596	2 US-08-194-088B-15	Sequence 15, Appl
C 11	78.4	11.9	10596	5 PCT-US93-04648-15	Sequence 15, Appl
C 12	75.4	11.4	3489	2 US-08-728-323A-1	Sequence 1, Appl
C 13	75.4	11.4	32207	2 US-08-770-379-20	Sequence 20, Appl
C 14	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 15	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 16	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 17	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 18	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 19	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 20	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 21	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 22	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 23	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 24	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 25	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 26	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 27	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl

C 28	57.6	8.7	2150	2 US-08-396-001-13	Sequence 13, Appl
C 29	57.6	8.7	2150	4 US-09-323-433A-13	Sequence 13, Appl
C 30	56.4	8.5	2340	3 US-09-022-983-4	Sequence 4, Appl
C 31	56.4	8.5	2477	4 US-09-490-692-3	Sequence 3, Appl
C 32	56.2	8.5	9636	1 US-08-323-170B-1	Sequence 1, Appl
C 33	55.6	8.4	16442	3 US-08-781-891-208	Sequence 208, App
C 34	54.8	8.3	397	3 US-09-253-691-3	Sequence 3, Appl
C 35	54.2	8.2	2277	1 US-08-676-967-2	Sequence 2, Appl
C 36	54.2	8.2	2277	1 US-08-676-967-2	Sequence 2, Appl
C 37	54.2	8.2	2277	1 US-08-676-967-2	Sequence 2, Appl
C 38	53.8	8.2	2301	1 US-08-306-691B-23	Sequence 23, Appl
C 39	53.8	8.2	2301	5 PCT-US93-06251-78	Sequence 78, Appl
C 40	53.6	8.1	19056	4 US-09-272-032-8	Sequence 8, Appl
C 41	53.2	8.1	4897	6 US-09-007-005-17	Sequence 17, Appl
C 42	53	8.0	289	4 US-09-244-796-17	Sequence 17, Appl
C 43	53	8.0	289	4 US-09-244-796-17	Sequence 17, Appl
C 44	53	8.0	1276	4 US-09-177-325-2	Sequence 2, Appl
C 45	53	8.0	1276	4 US-09-411-812A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22113-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-Fls
US-08-232-463-14

KW Dog: X-linked progressive retinal atrophy 2; XLPRA2; genetic marker;
 retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
 KW Miniature Schnauzer; mutant; mutein; ss.
 OS Canis familiaris.
 XX Synthetic.
 XX Key
 FT CDS
 FT 1..1149
 FT /tag= a
 FT /product= "Canine retinitis pigmentosa GTPase
 FT regulator mutant"
 FT /note= "CDS does not include start codon"
 FT /partial
 FT mutation
 FT replace (931..932, AGAG)
 FT /**tag= b
 FT /note= "This deletion results in the change of
 FT amino acids and ends in a premature stop codon"
 PN WO200138578-A1.
 XX 31-MAY-2001.
 XX 21-NOV-2000; 2000WO-US31940.
 XX 24-NOV-1999; 99US-0167365.
 XX (CORR) CORNELL RES FOUND INC.
 PA Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
 PI WPI: 2001-367707/38.
 XX P-PSDB; AAE02399.
 DR
 XX Identifying dogs with or carrying X-linked progressive retinal atrophy
 PT by detecting retinitis pigmentosa GTPase regulator gene mutation,
 PT useful when breeding Husky, Samoyed and Miniature Schnauzer
 XX
 PS Claim 75; Page 34-35; 88pp; English.
 XX
 CC The invention relates to a method for identifying dogs which are
 CC genetically normal, are carriers of, or are affected with X-linked
 CC progressive retinal atrophy (XLPRA), by testing a biological sample with
 CC genetic markers that co-segregate with a XLPRA gene locus. The invention
 CC also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
 CC mutants and their corresponding nucleic acid molecules. The mutated RPGR
 CC genes are responsible for the XLPRA in dogs. Methods are used to select
 CC dogs for breeding so that dogs carrying the mutated locus are eliminated
 CC from the breeding stock. The method particularly applies to Siberian
 CC Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
 CC the disease is X-linked. XLPRA1 type is identified in Siberian Huskies,
 CC and Samoyeds, while XLPRA2 type is identified in Miniature Schnauzers.
 CC The present CDNA sequence is the mutant open reading frame (ORF) 15
 CC encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant
 CC found in XLPRA2-affected dogs. This mutant is obtained by deleting
 CC 'GA' nucleotide bases from position 932 to 933 of the wild-type
 CC canine RPGR CDNA.
 XX
 SQ Sequence 2803 BP; 1028 A; 287 C; 939 G; 549 T; 0 other;

Query Match 12.4%; Score 81.6; DB 22; Length 2803;
 Best Local Similarity 49.2%; Pred. No. 4.6e-09;
 Matches 300; Conservative 0; Mismatches 304; Indels 6; Gaps 3;

OY 48 ggaatgggaagacagaagaagaatgcaaggagcatccagagagaagaagg 107
 DB 541 ggaagagggggaattggaagggaagggaaggaattagagaagaaggaggaagg 600
 OY 108 gaatgagggcgtgctcttcagggacgcgttgagcagttgaaaggcggagcg 167
 DB 601 gaagaagaagagggaattcggaagggggaaggaggaagggaagggaaggagag 660

OY 168 ggaagcgaagatcgaggagcagcagaaggagcagcgaggagcagaagcgcgagcgcg 227
 DB 661 aaagggagagcttgaaggggaggaaggagagcagaatctggaaaggaagggaagctggag 720
 OY 228 ggcggaagagcgccgcaagagcgaggagcccgcaaggaaagtctcgcacatcacccgac 287
 DB 721 gaagggagaagggaagtggagggaggaaggaggaagtggaagaagaaggagggaga 780
 OY 288 gatgagaagagactacagcagacaagtgaagccagccactggagtcgacgcccgctcg 347
 DB 781 ggaagagggggaagtggagg---aaggagaagcagaagaagagatctcagaagaagaaga 837
 OY 348 gccgcgcggagcggcttcaggtctggagcagcgccggaagccggtaaagaagaagaat 407
 DB 838 agggaaagtgaagaagaagaagag-aggggagaagcaacaagaagaatatagaaggaggga 896
 OY 408 ggaagaaggagacctgctgtcgcacttacaggagcatcagcagcagcgaaggaagcag 467
 DB 897 ggaagagggaggaagagaggaagaaagagaataatgaagaaggaggaaggaaggaag 956
 OY 468 ctgcgcgagctcctgcagcagaatcaggctcaggtctgaaggaagaaggaagga 527
 DB 957 caggg--gaaggggaggaaggaaggaagagagggcagcagcaggaagggaggggga 1014
 OY 528 ggaagagggaaggaggaaggaaggagcaccagtggaagatcaggggaaggaaggaagga 587
 DB 1015 agaagggggaaggaggaaggtagggaagaaagagggcgaaggaagaagaagaaggaagg 1074
 OY 588 ggaaggaggagccgcagcagcactctgaggagagcatcagcagctcgcgaagaagtaag 647
 DB 1075 ggaagagaaggaagaggaaggaagggaaggaagggaaggaagggaagggaagg 1134
 OY 648 tgaaggaga 657
 DB 1135 aggaaggga 1144

RESULT 14

ID AAD07130 standard; CDNA; 2805 BP.

AC AAD07130;
 DT 06-AUG-2001 (first entry)

DE Canine retinitis pigmentosa GTPase regulator (RPGR) cDNA.

KW Dog: X-linked progressive retinal atrophy; XLPRA; genetic marker;
 KW retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
 KW Miniature Schnauzer; ss.

OS Canis familiaris.

XX Key Location/Qualifiers
 FT CDS 1..1569
 FT /tag= a
 FT /product= "Canine retinitis pigmentosa GTPase regulator"
 FT /note= "CDS does not include start codon"
 FT /partial

PN WO200138578-A1.

PD 31-MAY-2001.

PE 21-NOV-2000; 2000WO-US31940.

PR 24-NOV-1999; 99US-0167365.

PA (CORR) CORNELL RES FOUND INC.

PI Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;

DR WPI: 2001-367707/38.

[illegible][illegible]

```

RESULT 11
AAI35486
ID AAI35486 standard; DNA; 1969 BP.
XX
XX
AC AAI35486;
XX
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #4172 used to measure gene expression in human placenta sample.
XX
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX
XX
KW genetic disorder; ss.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200157272-A2.
XX
XX
PD 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX
DR WPI; 2001-488897/53.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX
XX
PT analyzing gene expression in human placenta -
XX
XX
PS Claim 25; SEQ ID No 4172; 654bp; English.
XX
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
XX
CC The present sequence is one such probe. The probes are useful for
XX
CC producing a microarray for predicting, measuring and displaying gene
XX
CC expression in samples derived from human placenta. The probes are useful

```


DR WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
PS Claim 4; Page 32; 57pp; English.
XX
XX The PITSURE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated. p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the PITSURE protein kinase IRES. The
CC IRES sequence and fragments of it may be used to induce cell cycle
CC dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.
XX
SQ Sequence 222 BP; 77 A; 36 C; 89 G; 20 T; 0 other;
XX
Query Match 33.6%; Score 222; DB 21; Length 222;
Best Local Similarity 100.0%; Pred. No. 4.4e-40;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 439 gacatcagcgacagcgaggaagaccagctcgccgagctcctcgtcagcagaatcagc 498
DB 1 gacatcagcgacagcgaggaagaccagctcgccgagctcctcgtcagcagaatcagc 60
QY 499 tcaggtcttgaggaaag 558
DB 61 tcaggtcttgaggaaag 120
QY 559 gaagaaatcag 618
DB 121 gaagaaatcag 180
QY 619 gcatcagagcagctcgtccgaaagaagtaagtgaagaagaatg 660
DB 181 gcatcagagcagctcgtccgaaagaagtaagtgaagaagaatg 222
RESULT 5
AAA73710
ID AAA73710 standard; RNA; 222 BP.
XX
XX AAA73710;
AC
XX
XX 07-DEC-2000 (first entry)
DT
XX
DE RNA of human PITSURE protein kinase gene internal ribosome entry site.
XX
XX IRES; internal ribosome entry site; PITSURE; human; gene therapy;
KM cancer; restenosis; p58; p110; protein kinase; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH misc-signal 1..222
FT /tag= a
FT /label= IRES
FT /note= "internal ribosome entry site"
PN
XX WO200044896-A1.
XX
XX 03-AUG-2000.
PD
XX
XX 26-JAN-2000; 2000MO-EP00643.
PF
XX
XX 26-JAN-1999; 99EP-0200216.
PR
XX
XX (VLAAS-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
PA
XX

PI Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
DR
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
PS Disclosure; Page 32; 57pp; English.
XX
XX The PITSURE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated. p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the corresponding RNA of the PITSURE
CC IRES described in AAA73710. The IRES sequence and fragments of it may be
CC used to induce cell cycle dependent initiation of translation in
CC eukaryotic cells. Vectors containing the IRES may be used for the
CC preparation of compositions for the treatment of and/or prevention of
CC a disease by gene therapy. Such diseases may be cancer and restenosis.
XX
SQ Sequence 222 BP; 77 A; 36 C; 89 G; 20 U; 0 other;
XX
Query Match 33.6%; Score 222; DB 21; Length 222;
Best Local Similarity 91.0%; Pred. No. 4.4e-40;
Matches 202; Conservative 20; Mismatches 0; Indels 0; Gaps 0;
QY 439 gacatcagcgacagcgaggaagaccagctcgccgagctcctcgtcagcagaatcagc 498
DB 1 gacatcagcgacagcgaggaagaccagctcgccgagctcctcgtcagcagaatcagc 60
QY 499 tcaggtcttgaggaaag 558
DB 61 ucagguucugag 120
QY 559 gaagaaatcag 618
DB 121 gaagaaatcag 180
QY 619 gcatcagagcagctcgtccgaaagaagtaagtgaagaagaatg 660
DB 181 gcatcagagcagctcgtccgaaagaagtaagtgaagaagaatg 222
RESULT 6
AAA73715
ID AAA73715 standard; DNA; 67 BP.
XX
XX AAA73715;
AC
XX
XX 07-DEC-2000 (first entry)
DT
XX
DE Functional fragment of human PITSURE protein kinase gene IRES.
XX
XX IRES; internal ribosome entry site; PITSURE; human; gene therapy;
KM cancer; restenosis; p58; p110; protein kinase; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH misc-signal 1..222
FT /tag= a
FT /label= IRES
FT /note= "internal ribosome entry site"
PN
XX WO200044896-A1.
XX
XX 03-AUG-2000.
PD
XX
XX 26-JAN-2000; 2000MO-EP00643.
PF
XX
XX 26-JAN-1999; 99EP-0200216.
PR
XX
XX (VLAAS-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
PA
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
DR
XX

Db 709 cgcacaggagcgggagcccgcaagggaagtgtctgcacatcacccgaacgatgagagagac 768
Qy 301 tacagcgacaaagtgaagaagccagcacttgatctgcagcccgctcgccgcgcgagag 360
Db 769 tacagcgacaaagtgaagaagccagcacttgatctgcagcccgctcgccgcgcgagag 828
Qy 361 cgggttcgaatttgagagacggccggaagccagtaaaagaagaataatggaagaagagac 420
Db 829 cgggttcgaatttgagagacggccggaagccagtaaaagaagaataatggaagaagagac 888
Qy 421 ctgctgtccgacttacagacatcagcgacagcgagaggaagaccagctcgccgagtc 480
Db 889 ctgctgtccgacttacagacatcagcgacagcgagaggaagaccagctcgccgagtc 948
Qy 481 tcgtccacagaatcaggtctcaggttcgaggaagaagagagagagagagagagagag 540
Db 949 tcgtccacagaatcaggtctcaggttcgaggaagaagagagagagagagagagagag 1008
Qy 541 gaggaagggagcaccagatgaagaatcagagagaggaagagagagagagagagagac 600
Db 1009 gaggaagggagcaccagatgaagaatcagagagaggaagagagagagagagagac 1068
Qy 601 ggcagcactctgagagcgacatcagagcagctctgcgaagaagtaagtgaggaagaa 660
Db 1069 ggcagcactctgagagcgacatcagagcagctctgcgaagaagtaagtgaggaagaa 1128

RESULT 3
AAA73713 standard; DNA; 468 BP.

AC AA73713;

DT 07-DEC-2000 (first entry)

DE 5' extension of PITSLRE protein kinase internal ribosome entry site.

XX IRES: internal ribosome entry site; PITSLRE: human; gene therapy;

KW cancer; restenosis; p58; p110; protein kinase; ds.

XX Homo sapiens.

XX WO200044896-A1.

XX 03-AUG-2000.

XX 26-JAN-2000; 2000WO-EP00643.

XX 26-JAN-1999; 99EP-0200216.

XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Cornelis S, Beyaert R;

XX WPI: 2000-499331/44.

XX Nucleic acids encoding internal ribosome entry sequences useful for

PT directing protein expression in gene therapy procedures -

XX Claim 9; Page 33; 57pp; English.

XX The PITSLRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence contains a 5' fragment of the PITSLRE
CC protein kinase IRES and nucleotides upstream of the IRES. The IRES
CC sequence and fragments of it may be used to induce cell cycle
CC dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.

SQ Sequence 468 BP; 144 A; 100 C; 182 G; 42 T; 0 other;

Query Match 70.9%; Score 468; DB 21; Length 468;

Best Local Similarity 100.0%; Pred. No. 5.8e-94; Mismatches 0; Gaps 0;

Matches 468; Conservative 0; Indels 0;

Qy 1 caggaagcgaggaagacatcgagaagacagagataaagctcgccgggaatggaaga 60
Db 1 caggaagcgaggaagacatcgagaagacagagataaagctcgccgggaatggaaga 60
Qy 61 caggaaggaaggaatggaaggaagcattccaggaaggaaggaaggaatgagcgtg 120
Db 61 caggaaggaaggaatggaaggaagcattccaggaaggaaggaaggaatgagcgtg 120
Qy 121 tgcctcttaaggagccgcttgagcaggttagaagaagcggaagcgagagcgagagtg 180
Db 121 tgcctcttaaggagccgcttgagcaggttagaagaagcggaagcgagagcgagagtg 180
Qy 181 cgggaagcagcagaagagcagcgagcagaagagcgagcgagcgagcgagagcgag 240
Db 181 cgggaagcagcagaagagcagcgagcagaagagcgagcgagcgagcgagagcgag 240
Qy 241 cgcgaagagcggaagcgccgcaggggaagtgtctgcacatcacccgaacgagtgagagagac 300
Db 241 cgcgaagagcggaagcgccgcaggggaagtgtctgcacatcacccgaacgagtgagagagac 300
Qy 301 tacagcgacaaagtgaagaagccagcacttgatctgcagcccgctcgccgcgcgagag 360
Db 301 tacagcgacaaagtgaagaagccagcacttgatctgcagcccgctcgccgcgcgagag 360
Qy 361 cgggttcgaatttgagagacggccggaagccagtaaaagaagaataatggaagaagagac 420
Db 361 cgggttcgaatttgagagacggccggaagccagtaaaagaagaataatggaagaagagac 420
Qy 421 ctgctgtccgacttacagacatcagcgacagcgagaggaaggaagaccagc 468
Db 421 ctgctgtccgacttacagacatcagcgacagcgagaggaaggaagaccagc 468

RESULT 4

AAA73709 standard; DNA; 222 BP.

AC AA73709;

DT 07-DEC-2000 (first entry)

DE Human PITSLRE protein kinase gene internal ribosome entry site.

XX IRES: internal ribosome entry site; PITSLRE: human; gene therapy;

KW cancer; restenosis; p58; p110; protein kinase; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT misc_signal 1..222

FT /tag= a

FT /label= IRES

FT /note= "Internal ribosome entry site"

XX WO200044896-A1.

XX 03-AUG-2000.

XX 26-JAN-2000; 2000WO-EP00643.

XX 26-JAN-1999; 99EP-0200216.

XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Cornelis S, Beyaert R;

XX

XX Nucleic acids encoding internal ribosome entry sequences useful for
 PT directing protein expression in gene therapy procedures -
 XX
 PS Claim 10; Page 34; 57pp; English.

CC The PITSURE protein kinase gene can be translated to give two isoforms,
 CC p110 and p58. Transcription of p58 mRNA was found to start at an
 CC internal ribosome entry site (IRES). The IRES element was found to be
 CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
 CC cycle. The present sequence contains the PITSURE protein kinase
 CC IRES and upstream nucleotides. The IRES sequence and fragments of it
 CC may be used to induce cell cycle dependent initiation of translation in
 CC eukaryotic cells. Vectors containing the IRES may be used for the
 CC preparation of compositions for the treatment of and/or prevention of
 CC a disease by gene therapy. Such diseases may be cancer and restenosis.

XX Sequence 660 BP; 210 A; 128 C; 261 G; 61 T; 0 other;

Query Match 100.0%; Score 660; DB 21; Length 660;

Best Local Similarity 100.0%; Pred. No. 4.8e-136; Mismatches 0; Indels 0; Gaps 0;

Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caggaacgtcggaaacagacatcgagaagaagaagaataagctcgccgggaatgggaaga 60

DB 1 caggaacgtcggaaacagacatcgagaagaagaagaataagctcgccgggaatgggaaga 60

QY 61 caggaagaagaagaagaatgggaagaagaagaatggaagaagaagaagaagaagaaga 120

DB 61 caggaagaagaagaagaatgggaagaagaagaatggaagaagaagaagaagaagaaga 120

QY 121 tgcctcttcaggagaccgcttggaagcagttagaagaagcggagcggagcgaagctg 180

DB 121 tgcctcttcaggagaccgcttggaagcagttagaagaagcggagcggagcgaagctg 180

QY 181 cgggagcagcagaagaagcagcggagcagaagaagcggagcggagcggagcggagcgg 240

DB 181 cgggagcagcagaagaagcagcggagcagaagaagcggagcggagcggagcggagcgg 240

QY 241 cgcgaagagcggagcggagcggagcggagcggagcggagcggagcggagcggagcgg 300

DB 241 cgcgaagagcggagcggagcggagcggagcggagcggagcggagcggagcggagcgg 300

QY 301 tacagcgacaagaagtgaagccagccactggaagtcgcagcccgcccgcccgccggggag 360

DB 301 tacagcgacaagaagtgaagccagccactggaagtcgcagcccgcccgcccgccggggag 360

QY 361 cgggttcgagtcgggaagcggccgggaagccagtaaaagaagaagaatgggaagaagac 420

DB 361 cgggttcgagtcgggaagcggccgggaagccagtaaaagaagaagaatgggaagaagac 420

QY 421 cgcgttcgcgacttaacagacatcagcgacgaggaaggaaggaacacgagccgagccag 480

DB 421 cgcgttcgcgacttaacagacatcagcgacgaggaaggaaggaacacgagccgagccag 480

QY 481 tcgttcgacgaatcagggcccaagttctcgaggaagaagaagaaggaagaagaagagag 540

DB 481 tcgttcgacgaatcagggcccaagttctcgaggaagaagaagaaggaagaagaagagag 540

QY 541 gaggaaggaagcagcagatgaagaatcagaaggaagaagaagaagaagaagaagac 600

DB 541 gaggaaggaagcagcagatgaagaatcagaaggaagaagaagaagaagaagaagac 600

QY 601 ggaggaacactcgaggaagcagcagcttcgcgaagaagaataagtgagaagaagaatg 660

DB 601 ggaggaacactcgaggaagcagcagcttcgcgaagaagaataagtgagaagaagaatg 660

RESULT 2
 AAA73711
 ID AAA73711 standard; DNA; 2471 BP.

AC AAA73711;
 XX
 DT 07-DEC-2000 (first entry)
 XX

DE Human PITSURE protein kinase isoform alpha2-2.

KW IRES; internal ribosome entry site; PITSURE; human; gene therapy;
 KM cancer; restenosis; p58; p110; protein kinase; ds.

OS Homo sapiens.

PH Key Location/Qualifiers

FT msc_signal 907..1128

FT /*tag= a

FT /*tag= IRES

FT /note= "Internal ribosome entry site"

FN W0200044896-A1.

PD 03-AUG-2000.

XX 26-JAN-2000; 2000MO-BP00643.

XX 26-JAN-1999; 99EP-0200216.

XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Cornelis S, Beyaert R;

XX WPI; 2000-499331/44.

XX Nucleic acids encoding internal ribosome entry sequences useful for

XX directing protein expression in gene therapy procedures -

XX Disclosure; Page 32-33; 57pp; English.

XX The PITSURE protein kinase gene can be translated to give two isoforms,

XX p110 and p58. Transcription of p58 mRNA was found to start at an

XX internal ribosome entry site (IRES). The IRES element was found to be

XX cell cycle regulated, p58 is produced during the G2/M stage of the cell

XX cycle. The present sequence is the human PITSURE protein kinase isoform

XX alpha2-2. This sequence contains the IRES. The IRES sequence and

XX fragments of it may be used to induce cell cycle dependent initiation

XX of translation in eukaryotic cells. Vectors containing the IRES may be

XX used for the preparation of compositions for the treatment of and/or

XX prevention of a disease by gene therapy. Such diseases may be cancer

XX and restenosis.

XX Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other;

Query Match 100.0%; Score 660; DB 21; Length 2471;
 Best Local Similarity 100.0%; Pred. No. 6.4e-136;
 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caggaacgtcggaaacagacatcgagaagaagaagaataagctcgccgggaatgggaaga 60

DB 469 caggaacgtcggaaacagacatcgagaagaagaagaataagctcgccgggaatgggaaga 528

QY 61 caggaagaagaagaatgggaagaagaagaatggaagaagaagaagaagaagaagaaga 120

DB 529 caggaagaagaagaatgggaagaagaagaatggaagaagaagaagaagaagaagaaga 568

QY 121 tgcctcttcaggagaccgcttggaagcagttagaagaagcggagcggagcgaagatg 180

DB 589 tgcctcttcaggagaccgcttggaagcagttagaagaagcggagcggagcgaagatg 648

QY 181 cgggagcagcagaagaagcagcggagcagaagaagcggagcggagcggagcggagcgg 240

DB 649 cgggagcagcagaagaagcagcggagcagaagaagcggagcggagcggagcggagcgg 708

QY 241 cgcgaagagcggagcggagcggagcggagcggagcggagcggagcggagcggagcgg 300

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 19:03:42 ; Search time 355.05 Seconds
(without alignments)
1593.676 Million cell updates/sec

Title: US-09-915-060-6
Perfect score: 660
Sequence: 1 cagcaagctcggaacgaca.....aagtaagtgaggaagaatg 660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: N.Geneseq_1101.*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	660	100.0	660	21	AAA73714 Human PITSIRE prot
2	660	100.0	2471	21	AAA73711 Human PITSIRE prot
3	468	70.9	468	21	AAA73713 5' extension of PI
4	222	33.6	222	21	AAA73709 Human PITSIRE prot
5	222	33.6	222	21	AAA73710 RNA of human PITSIR
6	87	13.2	87	21	AAA73715 Functional fragmen
7	87	13.2	575	22	AA123315 Probe #13248 for g
8	87	13.2	575	22	AA148636 Probe #17322 used t
9	87	13.2	575	22	AA108956 Probe #8947 used t
10	87	13.2	1969	22	AA114105 Probe #4038 for ge
11	87	13.2	1969	22	AA135486 Probe #4172 used t

12	87	13.2	1969	22	AA103958
13	81.6	12.4	2803	22	AAAD07132
14	81.2	12.3	2805	22	AAAD07130
15	79.6	12.1	2800	22	AAAD07131
16	78.6	11.9	1925	20	AAAX90924
17	78.4	11.9	1926	21	AAAS50254
18	78.4	11.9	1926	22	AAAF82902
19	78.4	11.9	2580	21	AAAF7454
20	78.4	11.9	5452	20	AAAX90923
21	78.4	11.9	8705	20	AAZ23778
22	78.4	11.9	9600	19	AAV21683
23	78.4	11.9	10380	20	AAZ22248
24	78.4	11.9	10596	14	AAO51731
25	78.4	11.9	10596	17	AAI40348
26	78.4	11.9	10596	20	AAI5650
27	78.4	11.9	16080	21	AAAS5553
28	75.4	11.4	3489	21	AAAS30290
29	75.4	11.4	3489	22	AAAF82901
30	75.4	11.4	32207	20	AAV73805
31	75.4	11.4	137507	19	AAV19941
32	75.2	11.4	795	19	AAV55830
33	75.2	11.4	799	19	AAV55831
34	66.8	10.1	49999	20	AAZ23891
35	66.8	10.1	49999	20	AAZ23896
36	65.4	9.9	49999	20	AAZ23901
37	65	9.8	15672	12	AAO10613
38	64.6	9.8	2643	14	AAO39212
39	64.6	9.8	3717	21	AAAF6460
40	64.2	9.7	470	22	AAI11522
41	64.2	9.7	470	22	AAI3875
42	64.2	9.7	470	22	AAI02430
43	63.6	9.6	726	22	AAI22214
44	63.6	9.6	726	22	AAI47511
45	63.6	9.6	726	22	AAI07913

ALIGNMENTS

RESULT 1
ID AAA73714 standard; DNA; 660 BP.
AC AAA73714;
XX
DT 07-DEC-2000 (first entry)
XX
DE Human PITSIRE protein kinase gene IRES and upstream sequence.
XX
XX IRES: internal ribosome entry site; PITSIRE: human; gene therapy;
KW cancer; retinosis; p58; p110; protein kinase; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_signal 439..660
FT /*tag= a
FT /label= IRES
FT /note= "Internal ribosome entry site"
XX
XX WO20044896-A1.
XX
PD 03-AUG-2000.
XX
PF 26-JAN-2000; 2000WO-EP00643.
XX
PR 26-JAN-1999; 99EP-0200216.
XX
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Cornelis S, Beyaert R;
XX
DR WPI; 2000-499331/44.

Probe #3949 used t
Canine retinitis p
Canine retinitis p
Canine retinitis p
Epstein Barr Virus
Epstein Barr Virus
EBV tethering prot
Nucleotide sequenc
Anti-sense strand
Vector pshuttle DN
Vector plasmid PCM
Nucleotide sequenc
Plasmid pcISBON f
Plasmid pcISBON f
Nucleotide sequenc
DNA clone pCK C1.
Kapost's sarcoma-a
Nucleotide sequenc
KSHV LTR DNA (nucl
KSHV long unique c
FLGA insert stabl
Nucleotide sequenc
Murine LOBO genom
Murine LOBO homolo
Human LOBO homolo
Rianodin receptor
CENP-B CDNA. Homo
DNA encoding cent
Probe #2455 for ge
Probe #2561 used t
Probe #2421 used t
Probe #12147 for g
Probe #16197 used t
Probe #7904 used t

gene . /map="1p36.3"
1. .2533
/gene="CDC2L1"
CDS 261. .2507
/gene="CDC2L1"
/note="corresponding genomic sequence deposited as Genbank
Accession Numbers AF080678-AF080688, AF092429, and
AF092430"
/product="Pif1S1 protein kinase alpha SV10 isoform"
/codon_start=1
/db_xref="GI:3850312"
/protein_id="AAC72081.1"
/translation="MSQSDRDSKRDLSLEGELRDHRWITIRNSPYRSDMEDRGE
EDDSLAIRPQOAKREKRVHHRKDEKRRHSHSAAGGHAHAKERERERKER
HREEDKARREMREREMARSHSRERDLERERERERERERERERERERER
ERAEER
KEKMEERDLSDLOIDSEKRTSSAESASGSGSEPEERENHLLVPSRTRDS
EPEER
GSEER
CLNRIEETGYVYVYRAKDKTDEIVALKRLKMEKEEFPITSLREINTILKAHPNI
VYVREIVGSSNDKLYIVNRYVEHDLKSLMETMOPFLPEVKTLMIOLRGVKHLHD
NMLHRDLKTSNLLSHAGILKVGDFGLAREXSPKATPVVITLWYAPELLGAK
EYSTAVDMSSVCIFEGELTOKPLFPGKSEIDQINKFKDILGTPSEKIMPGYSELPAV
KMTFSEHPYNNLRKREFGALLSDOGFDLMNKFLIYFPGRIISAEDGLKHEYRETPLP
IDPSMPTTPAKSEQQRVVRKGTSPRPPEGGLGYSQLGDDDLKETGFHLTTINOGASAA
GPEFSLK"

BASE COUNT 715 a 615 c 826 g 377 t
ORIGIN

Query Match 87.4%; Score 577; DB 9; Length 2533;
Best Local Similarity 94.6%; Pred. No. 2.7e-97;
Matches 633; Conservative 0; Mismatches 0; Indels 36; Gaps 2;

QY 1 CACGAACGTCGGAACGACATCGAGAAGAACGAATCAAGTCGCGGGAATGGAAGA 60
DB 549 CACGAACGTCGGAACGACATCGAGAAGAACGAATCAAGTCGCGGGAATGGAAGA 608
QY 61 CAGAAGAAGAGGAATGCGAAGGAGCATCCAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 609 CAGAAGAAGAGGAATGCGAAGGAGCATCCAGAGAGAGAGAGAGAGAGAGAGAG 650
QY 121 TGCCTCTCAGGACCGCTTGAGCACTGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 651 -----AGGACCGCTTGAGCACTGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 701
QY 181 CCGGAGCAGCAG 240
DB 702 CCGGAGCAGCAG 761
QY 241 CCGAAG 300
DB 762 CCGAAG 821
QY 301 TACAGCAACAAGTGAAGCCAGCCACTGAGTGCAGCCCGCTGCGCGCGGAGAG 360
DB 822 TACAGCAACAAGTGAAGCCAGCCACTGAGTGCAGCCCGCTGCGCGCGGAGAG 881
QY 361 CCGTTCGAGTTGGAGAGCGCCGAGAGCCAGTAAAGAGAGAGAGAGAGAGAGAGAG 420
DB 882 CCGTTCGAGTTGGAGAGCGCCGAGAGCCAGTAAAGAGAGAGAGAGAGAGAGAGAG 941
QY 421 CTGCTTCGACTTACAGGACATCAGCGACCGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 942 CTGCTTCGACTTACAGGACATCAGCGACCGAGAGAGAGAGAGAGAGAGAGAGAG 1001
QY 481 TCGTCAGCAAGATCAAGCTCAGGTTCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 1002 TCGTCAGCAAGATCAAGCTCAGGTTCTGAGCAAGAGAGAGAGAGAGAGAGAGAG 1061
QY 541 GAGGAAGGAG 591
DB 1062 GAGGAAGGAG 1121

QY 592 GAGGAAGCAGCAGCACTCTGAGGAGGATCAGAGCAAGTCTGCCGAAGAGAGAGAGAG 651
DB 1122 GAGGAAGCAGCAGCACTCTGAGGAGGATCAGAGCAAGTCTGCCGAAGAGAGAGAGAG 1181
QY 652 GAAGAAGATG 660
DB 1182 GAAGAAGATG 1190

Search completed: December 4, 2001, 18:54:48
Job time: 9184 sec

Db	576	CAGAAAGACAAGGAAATGTGGCAAGAGGAGCATTTCCAGGAGAGAA-----	61/
QY	121	tgcccttcaggagaccgcttgtagaacgttagaagaagagcgggagcggaagcgcaagatg	180
Db	618	-----AGGAGCCGCTTTGGAGCAGTTTAAGAAAGAACGGGAGCGGACGCAAGATG	668
QY	181	cggagcagcagaagagcaagcggagagcagaagaagagcgcgagcgggcggaagagcgg	240
Db	669	CGGAGCGAGCGAGAGGAGCAGCGGGAGCAGAAAGAGCGCGAGCGGGCGGAGAGCGG	728
QY	241	cgcaagagcggggagggcccgaggggaadtgtctgcacatacggaaagataagaaagagc	300
Db	729	CGCAAGAGCGGGAGGCCCGCAGAGGAAGTGTGTGCATCACCGAACGATGAGAGAGAC	788
QY	301	tacacgcacaaagtgaagagccagccactgaggtcgcgaagcccgctcgagcgcgcgaggag	360
Db	789	TACACGCACAAGTAGTAAGAACCGCACACTGTGATCGCACCCGCTCGCGCCGCGGGAG	848
QY	361	cggcttcgagttgggagacagcgccggagaccagtaagaagaagaaatgaaagaaagagac	420
Db	849	CGGTTTCGAGTTGGGAGAGCGCGCGAGAGCCAGTAAAGAGAGAAATGAGAAAGGAGAC	908
QY	421	ctgcgtccgacttacagagacatacgcagacagcgagagaaagaccagctcggcgagtcgc	480
Db	909	CTGCTGTCCGACTTACAGGACATCAGCCAGCAGAGAGAGAACACAGCTCGGCGAGTCC	968
QY	481	tcgtcagcagaatcaggtcaggtctcgtctgaagaaagaagagagagagagagagag	540
Db	969	TGCTGACGACGATCAGGCTCAGGTTCTGAGGAAGAGAGAGAGAGAGAGAGAGAGAG	1028
QY	541	gaggaagagagagagacgaatggaagaatca-----gagggggagagagagagagagag	591
Db	1029	GAGGAAGGAGAGCACCAGTGTGAATATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1088
QY	592	gagggagaccgggagcaactctgaaggagagcatcagagcagctcgcggaagaaatgag	651
Db	1089	GAGGAGACCGGACGACACTGTGAGAGCATCAGAGCAGTGTGCCGAGAAGTAGTAGAG	1148
QY	652	gaaagaatg 660	
Db	1149	GAAGAAATG 1157	
RESULT 15			
AF067516			
LOCUS	AF067516	2533 bp	mRNA
DEFINITION	Homo sapiens PITSIRE protein kinase alpha SV10 isoform (CDC2L1)		PRI
ACCESSION	AF067516		
VERSION	AF067516.1		
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			

QY	121	tgctcttcaggagccgccttgagcaagtagaagaagccggagccggagccgaagatg	180
Db	571	-----AGGACCGCGTTGGAGCGAGTTAGAAAGAAAGCGGAGCGGACCGCAAGATG	621
QY	181	cggagacagcagaagagagcagcgaggacagaagaagcgcagacggcgcgagagacgg	240
Db	622	CGGAGCAGCACAAGAGSAGCAGCGGGAGCAGAAAGAGCGCGAGCGGCGAGAGCGG	681
QY	241	cgcagaagagcgggagggcccgcaagggaagtgtctgcacatcaaccgaagatgagagagac	300
Db	682	CGAAGAGACGGGAGAGCGCCGCAAGGAAGTGTCTGCACATCACCGAACATGAGAGAGAC	741
QY	301	tacagcgacaagaatgaaagccagccacatggagatgcacgccgcctcgcccgcgagag	360
Db	742	TACAGGACAAAGTGAAAGCCAGCCACTGAGATGCGCACCCGCTCGGCGCCCGGAG	801
QY	361	cggcttcgagatctggagagagcgccggaagccagtaaaagaagaaaatggaagaagagac	420
Db	802	CGGTCGAGATTGGGAGACGCGCCGAGAGCCAGTAAAGAGAAAGCAAAATGCAAGAAAGGAC	861
QY	421	ctgcgtccgaattcacaagacatcgacagacgcagagagagagagacacagcttgcgagatcc	480
Db	862	CGGCTGTCGCAATTACAGGACACTACGCCACAGAGAGAGAAAGACCAGCTCGGCGAGTCC	921
QY	481	tcgtacagcgaatcagctcagctcaggtcttgaagaaagaagagagagagagagagagag	540
Db	922	TGCTAGCAGAAATCAAGGCTCAGGTTCTGAGAGAAAGAGAGAGAGAGAGAGAGAGAG	981
QY	541	ggaggaagagagacacagcagtgaaagaatcagaagagagagagagagagagagagagagac	600
Db	982	GAGGAAGGAGAGCACCAGTGAAGAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAAC	1041
QY	601	ggcagcaactctgaagaggaatcagagagagctctgcgaagaagatgaatgagagaaatg	660
Db	1042	GCGACCACTCTGAGAGAGCATCAGAGCAGCTGCGCAAGAACTAAGTAGAGAGAAATG	1101
RESULT 11			
LOCUS	HSU07704		
DEFINITION	Human protein kinase PITSLRE isoform PBETA21 (p58CDC2L1) mRNA,		
ACCESSION	U07704		
VERSION	U07704.1		
KEYWORDS	GI:507426		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 2471)		
TITLE	Xiang,J., Ishii,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.		
	Molecular cloning and expression of alternatively spliced PITSLRE		
	protein kinase isoforms		
JOURNAL	J. Biol. Chem. 269, 15786-15794 (1994)		
MEDLINE	94253170		
REFERENCE	2 (bases 1 to 2471)		
AUTHORS	Kidd,V.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-MAR-1994) Vincent J. Kidd, St. Jude Children's Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St., Memphis, TN 38101, USA		
FEATURES	Location/Qualifiers		
source	1..2471		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="1p36-2"		
	/cell_line="Hela S3"		
	/cell_type="epitheloid"		
	/tissue_type="cervix"		
	/clone_lib="Hela cDNA library"		
	/dev_stage="adult"		

QY	Db	434	CACGACGCTCGGAACGACATCGGAAGAACACAGATTAAGCTGCCCGGAATGGGAAGA	493
QY	61	cagaagaagaagygaaatgycagaaggaagcatccagsgagaagaaggggaatgatlgtgcgtg	120	
Db	494	CAGAAGAAGAAGGGAATGTGCGAAGGAGCAATTCCTCGAGAGAGNA	535	
QY	121	tgctcttcagggccgcttggagcagtttgaagaagcgggagcgggagcgaagatg	180	
Db	536	-----AGGGACCCCTTGGACACTTAAAGAGAAAGCGGAGCGGAGCGCAAGATG	586	
QY	181	cgaggagcagcagaagggagcagcgggagcagaagaagcgcgagcggcgggcggaagacgg	240	
Db	587	CGGAGACGACGAAAGAGACACCGGAGCAGAGAAGAGCGCGAGCGGGCGGAGAGCGG	646	
QY	241	cgcagaagcggggaagcccgcaaggaaagtgtctcaatcacccgaacgaatgagagagac	300	
Db	647	CGCAAGGAGCGGAGAGCCCGCAGGGAAGTGTCTGCACATCACCGAAGCATGAGAGAGAC	706	
QY	301	tacagcgacaagtgtgaaagccagccactcgggtgtgcgccccgcttcggccgcccggag	360	
Db	707	TACAGCGCAAAAGTGAAGAACCCAGCCACTGTGAGTGTGCACACCCGCTCGGCGCCGGGAG	766	
QY	361	cgatttcaggttgggagacgcgcgggaagccagtaaaagaagaagaatgaaagaagagac	420	
Db	767	CGGTTCCGAGTTGGGAGACGCGCGGAGACCCACTAATAAGAGAAGAAATGTGAAGAAAGGAC	826	
QY	421	ctgctgtccgaacttacaagagcatcagcgacacgcgagaggaagaccagctcggccgagtc	480	
Db	827	CTGCTGTCCGACTTACAGAGCATCAGCGACAGCGAGAGAGAACCCAGCTCGGCGCCAGTCC	886	
QY	481	tcgtcagcagaatacagagctcaggttcttgaagaagaagggaggaggaggaagagagag	540	
Db	887	TCTGTACGAGATCATCGGCTCAGGTTCTTAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG	946	
QY	541	gaggaaagggaacacacagcgaagaatcagaagagagaagaaggaagagagagagagacc	600	
Db	947	GAGGAGGAGGACACACAGTGAAGAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAC	1006	
QY	601	ggcagcaactctbgagagagcatcagagcagctctgcgaagaagaatgagtgaaggaagaa	660	
Db	1007	GCGACCAACTCTGAGAGAGGATCAGAGCAGTCTCCCGAAGAAAGTAAGTGAGGAAAGAAATG	1066	
RESULT	9			
AF067519				
LOCUS	2439 bp	mrna	PRI	07-NOV-1998
DEFINITION	Homo sapiens PITSRLRE protein kinase beta SV1 isoform (CDC2L2) mRNA,			
ACCESSION	AF067519			
VERSION	AF067519.1	GI:3850317		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 2439)			
JOURNAL	Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,			
REFERENCE	Ambros,P., and Kidd,V.J.			
AUTHORS	Duplication of a genomic region containing the Cdc2L1-2 and			
TITLE	MMP21-22 genes on human chromosome 1p36.3 and their linkage to D12z			
JOURNAL	Genome Res. 8 (9), 929-939 (1998)			
REFERENCE	98424414			
AUTHORS	2 (bases 1 to 2439)			
TITLE	Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,			
JOURNAL	Ambros,P., and Kidd,V.J.			
REFERENCE	Direct Submission			
AUTHORS	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's			
TITLE	Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA			
FEATURES	Location/Qualifiers			
SOURCE	1..2439			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="1"			

[illegible]

TITLE Ambros, P., and Kidd, V.J.
Direct Submission
JOURNAL Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES
source location/Qualifiers
1..2323
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2323
/gene="CDC2L2"
1104..2297
/gene="CDC2L2"
/note="corresponding genomic sequence deposited as Genbank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PITSLRE protein kinase beta SV8 isoform"
/protein_id="AAC72083.1"
/db_xref="GI:3850316"
/translation="MKNEKMTTSMLFQSHGSTETIPGRVKRKQKKMGCRSYEEFQCLN
RIEEGYGVYRAKDKTDEIVALKRKKEKEKEPEPTSLREINTILIAKQHPNIVTV
REIVYGSNDKIYIWMNVYHDLKSLMETMKRPPIRGKVTIMTGLKGVKTHDNMI
LHRDLKTSNLSLHAGILKVGDEGLAREVGSPLKATYTPVYTQWTRAPBLGAKETS
TAVDMASVCITCELTOKPPLPPGNSLIDQIKKVEKGTSEKTIWPAEYLPAKETS
TSEHFINLNKRFGLSLSDGFDLMNKFLEYTPGRIRISAEIDGLHFEYRETPIDP
SMFPTWPAKSEDOVRKRGTSPPRPEGGLGYSLDGDDEKTEGFLHTTNOGASAGPG
SLKRF"

gene
CDS

BASE COUNT 673 a 550 c 754 g 346 t
ORIGIN

Query Match 90.3%; Score 596; DB 9; Length 2323;
Best Local Similarity 95.9%; Pred. No. 8.4e-101;

Matches 633; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

1 cagaagcgtcggaaacagacatcgagaagaacagataaagctcgccggatgaggaaga 60
|||||
464 CAGGAACGTGGGAAACGACATCGAAGAAACAGATAAAGCTCCGCGGATGGGAAAGA 523
|||||
61 cagaagaagaagaatggaagagcattccagagagaagaaggagatgatgagcgtg 120
|||||
524 CAGAAGAAGAAGGAATGGCAAGGAGCATTCAGAGAGAA----- 565
|||||
121 tgcctcttaaggagccgcttgtagagcattagaagaagcggagagcgagagcgaatg 180
|||||
566 -----AGGACCGCTTGGAGCGATTGAAGAAAGCAAGCGAGCGGCAAGATG 616
|||||
181 cggagcagcagaagaagcggagcggagcgaagagcgagcggcgagcgaagagcg 240
|||||
617 CCGGACACACAGAGAGACAGCGGAGAGAGAGAGCGGAGCGGCGGAGAGAGCGG 676
|||||
241 cgaagaagcggagagcccgagagagatgtctgcacatcagcagaacgaatgagagagc 300
|||||
677 CGCAAGAGACGGGAGGCGCGAGGAGAGTGTGACATCACCGAAGCATGAGAGAGAGC 736
|||||
301 tacaagagaagaagtgaagaacccagcttgagtcgagccgctcgagcgcgagagag 360
|||||
737 TACAGCGACAAAGTGAAGACCACTGAGTCCGAGCCCTCGGCCCGCGCGGAG 796
|||||
361 cggcttcgaattggaagcggcggagacagcaaaagaagaagaatgagaagaagagc 420
|||||
797 CGGTTGAGTTGGAGAGCGCGGAGAACCAATGAAGAAAGAAATGAGAGAAAGGAGC 856
|||||
421 cgcgctccgacttaagagacatcagcgaagcgaagaggaagcagctcgagcgaatcc 480
|||||
857 CTGCTGTCCGACTTACAGACATCAGCGACACGAGAGAAACACCTCGCGCGAGTCC 916
|||||
481 tgcgcagagaatcagagccagcttctgaagagaagaaggagaggaagaagaagag 540
|||||
917 TCGTCAAGCAATCAAGCTCAAGCTTCTGAGGAAGAAGAGAGAGAAAGAGAGAG 976
|||||
541 gaggagagagacacagatgaagatcagagagagaggaaggaagagagagagagc 600
|||||

|||||
Db 977 GAGGAAGGAGACACAGTGAATCAGACAGGAGAGAGAGAGAGAGAGAGAC 1036
|||||
QY 601 ggaagcaactctgagagagcagcagcagcttcggaagaagtaagtgaagaagag 660
|||||
Db 1037 GCAGCAACTCGAGAGAGCATCAGAGCAGTCTGCCAAGAAAGTACGAGAGAAATG 1096
|||||

RESULT 8
AF067521

LOCUS AF067521 2409 bp mRNA PRI 07-NOV-1998
DEFINITION Homo sapiens PITSLRE protein kinase beta SV3 isoform (CDC2L2) mRNA,
complete cds.
AF067521
ACCESSION AF067521.1 GI:3850321
VERSION
KEYWORDS
SOURCE

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2409)

REFERENCE

Autors
Guturajan, R., Lahti, J.M., Grenet, J., Easton, J., Gruber, I.,
Ambros, P.F. and Kidd, V.J.
Duplication of a genomic region containing the Cdc2L2-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2
98424414
2 (bases 1 to 2409)

JOURNAL

MEDLINE

REFERENCE

AUTHORS

FEATURES

1..2409
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2409
/gene="CDC2L2"
80..2383
/gene="CDC2L2"
/note="corresponding genomic sequence deposited as Genbank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PITSLRE protein kinase beta SV3 isoform"
/protein_id="AAC72083.1"
/db_xref="GI:3850322"
/db_xref="GI:3850316"

gene
CDS

BASE COUNT 686 a 586 c 780 g 357 t
ORIGIN

Query Match 90.3%; Score 596; DB 9; Length 2409;
Best Local Similarity 95.9%; Pred. No. 8.3e-101;

Matches 633; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

1 cagaagcgtcggaaacagacatcgagaagaacagataaagctcgccggatgaggaaga 60
|||||

QY 421 ctgtcttccgacttaccagcaccatcagcagcagagagaaagaccagctcgccgagctcc 480
|||||
Db 671 CTGCTGTCCGACTTACAGGACATCAGCAGACGAGAGGAGACCGCTGGCCGAGTCC 730
QY 481 tcgtcaacaatcagctcaggtctctgagaaagagagagagagagagagagagag 540
|||||
Db 731 TCCTCAGCAATCAGGCTCAGGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 790
QY 541 gaagaaagagagacacagtgagaatacagagagagagagagagagagagagagacc 600
|||||
Db 791 GAGGAAGGAGGACACAGTGAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 850
QY 601 ggcagcaactctgagagagatcagagcagctctgccaaagtaagtgaagagaaatg 660
|||||
Db 851 GGCAGCACTCTGAGGAGCATCAGAGCAGTCTGCCAAGAGTAAGTGAAGAGAAATG 910

RESULT 6
AF067520 2436 bp mRNA PRI 07-NOV-1998
LOCUS AF067520 PRTSLRE protein kinase beta SV2 isoform (CDC2L2) mRNA,
DEFINITION complete cds.
ACCESSION AF067520.1 GI:3850319
VERSION AF067520
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2436)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to D122
JOURNAL Genome Res. 8 (9), 929-939 (1998)
MEDLINE 98424414
REFERENCE 2 (bases 1 to 2436)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES
source
location/Qualifiers
1..2436
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2436
/gene="CDC2L2"
80..2410
/gene="CDC2L2"
/note="corresponding genomic sequence deposited as GenBank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PRTSLRE protein kinase beta SV2 isoform"
/protein_id="GI:3850320"
/db_xref="GI:3850320"
/translation="MGDEKSKWKYKTLDELQKKRRKDEBEKAEIKRLKNSDRSK
RDSLEGEELDHCEITRNSPYRREDMSKDEEDSLAKFPQDSMEKHHNKK
EKREKWKHARVREHREKRRHREEDQKRRERARERREARERVSAAHRTMED
FDRLEQLERRRERERKREDOKEOREKREARERREARERVSAAHRTMED
YSDVKASHRSRPPRRPREFELGDRKPVKEKMERDLSDLDIDSEKRTSSA
ESSASGSGSEEEEEESEEGSTSESESESESESESESESESESESESESESE
EESE
EESE
ALPLTELKQELPKLPLALQCSRVEEFQCLNREEGTYGVYTRAKDKTDYALRL
KMEKEKGFITSLREINTILKAQHPNIVTRELIVGSSNDKITYVNIYEHDLKSLM
ETMKQPLPGEVTKLIMQLRGVKNLHDNMLIRDLKTSNLLSHAGILKVGFGSLAR
EYSGPLKATPVVVTOWYRAPPELLGAKESTAVDMSCVICIGELITOKPLPFGASE
IDQINKVFKEIGMPSEKIMPYSELVVKKMTSEHPNNLRRKFGALLSDQGFILMN
KFLTYFGRRTISADGKHEHRETPLPDPSMPMPMPASSEQGRVKGSPSPPEEG
LGYSQDDDDDKETGERHUTTTNGASAGGFSLKF"

BASE COUNT 690 a 591 c 789 g 366 t

ORIGIN
Query Match 99.5%; Score 656.8; DB 9; Length 2436;
Best Local Similarity 99.7%; Pred. No. 4.9e-112;
Matches 658; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 caccagactcgaacacgcatcgaagaacagatgaagctccgcggaattggaaga 60
|||||
Db 434 CACGAACGTCGAAACGACATCGAAGACAGATTAAGCTCGCGGGATGGAAAGA 493
QY 61 cagaagaagaagaaatgycagaaggagatccagaagaagaaggagagatgagcgctg 120
|||||
Db 494 CAGAAGAGAGGGAATGGCAAGGAGCATTCACAGAGAGAAAGGGGGAATGATGCTTT 553
QY 121 tgcctcttaaggagccgcttggagcafttagaagaagcgggagcggagcgaagattg 180
|||||
Db 554 TCCTCTTCAGGACCGCTTGAGCACTTAGAAAGAAAGCGGAGCGGAGCGAAGATG 613
QY 181 cggagagcagaagaagcagcagcggagcagaagagcgcgagcggcgagagcg 240
|||||
Db 614 CGGAGACACAG 673
QY 241 cgcagaagcggagggccgcgaggaagtgtctgcacatcacggaacgaatgagaagagc 300
|||||
Db 674 CGCAAGCAGCGGAGAGCGCCGCAAGCAAGTCTGCACATCACGGAACGATGAGAGAGAC 733
QY 301 taacagcagaagaagtgaagccagcagctgagctgcaagccgctccgcccgcggagag 360
|||||
Db 734 TACAGGACAAAGTGAACCCAGCCACTGGAGTCCGAGCCCTCGCCGCGCGGAG 793
QY 361 cggctcgaattggagagcgcgcggaagccagtaaaagaagaagaatggaagaaggagc 420
|||||
Db 794 CGGTTGAGATTGGGAGACGCGCGGAAGCCAGTAAAGAGAAATGAGAAAGAGGAC 853
QY 421 ctgtcttccgacttaccagacatcagcagcagcagagagagagagagagagagagc 480
|||||
Db 854 CTGCTTCGCACTTACAGGACATCAGCAGCAGCAGAGAGAGAACCAAGCTCGCCAGTCC 913
QY 481 tcgtcaacaatcagctcaggtctctgagaaagagagagagagagagagagagag 540
|||||
Db 914 TCCTCAGCAATCAGGCTCAGGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 973
QY 541 gaagaaagagacacagtgagaatacagagagagagagagagagagagagagacc 600
|||||
Db 974 GAGGAAGGAGGACACAGTGAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1033
QY 601 ggcagcaactctgagagagcagcagctctgccaaagtaagtgaagagaaatg 660
|||||
Db 1034 GGCAGCACTCTGAGGAGCATCAGAGCAGTCTGCCAAGAGTAAGTGAAGAGAAATG 1093

RESULT 7
AF067518 2323 bp mRNA PRI 07-NOV-1998
LOCUS AF067518 PRTSLRE protein kinase beta SV2 isoform (CDC2L2) mRNA,
DEFINITION complete cds.
ACCESSION AF067518.1 GI:3850315
VERSION AF067518
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2323)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to D122
JOURNAL Genome Res. 8 (9), 929-939 (1998)
MEDLINE 98424414
REFERENCE 2 (bases 1 to 2323)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,

LKMEKEGFPITSLREINTILKAQHNPITVREIVGSMNDKITVIMNVHEHDLKSL
METMKQPLFGEVTKLMTIOLRGVKNLHNDMLIHLRLKTSNLLSHAGILKQDFELA
REYSGPLKATPVVVTILWYRAPELLGAEKSEIVAVMWSVGLFEGELLQKPLFPKLS
EIDQINKVEFDLGTPESEKIMPGYSELPAVKMFPSEHPYNNIRKREGALLSPOGPDLM
NKLTYFPGRRISAEDGLKHEVRETPLEIDPSMFTPTAKESQOQVKKGISPRPEGS
GLQYSOLGDDDLKETETFHLTTNNOGASAGPESLKE"

BASE COUNT 697 a 601 c 809 g 364 t
ORIGIN

Query Match 100.0%; Score 660; DB 9; Length 2471;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caggaagcgtcgaaagacacatcgagaagacaagataaagctcgccgggaatgggaaga 60
|||||
DB 469 CACGAAGCTCGAAGACACATCGAAGAAAGACAGATTAACCTCGCCGGGATGGAAAGA 528
|||||
QY 61 cagaagaagaaagaaatgagcaagggagcaltccaagagaagaagggaatgagcgatg 120
|||||
DB 529 CAGAAGGAAAGGAAAGGAGGAGGAGCATTCAGAGAGAAAGGGGGAATGATGCGGTG 588
|||||
QY 121 tgcctcttcagagaccgcttgagacagtttagaagaagcggagcggagcggcaaatg 180
|||||
DB 589 TGCCTCTTACGGGACCCCTTGGAACATTTGAAAAGAGACGGAGCGGAGCCCAAGATG 648
|||||
QY 181 cgggaagcagagaagaaagcagcgaggaagaaagcagcgagcgagcgaggaagcg 240
|||||
DB 649 CGGAGAGCAGAGAGAGAGAGCGGAGCAGAAAGAGAGCGGAGCGGCGGAGGAGCGG 708
|||||
QY 241 cggcaagaagcggagagcccgacaggaagtgctgcacatccagaaacgatgagaagagc 300
|||||
DB 709 CGCAAGAGAGCGGAGGCGCCGAGGAGAAAGTGTCTGCACATCACGGAACGATGAGAGAGGAGC 768
|||||
QY 301 taacagcagacaagatgaaagccagccacgtgagtcgagccgctcgccgagcgagga 360
|||||
DB 769 TACAGCGACAAAGTGAAGAGCCAGCCACTGGAGTCGAGCCCGCTCGCCGCGGAG 828
|||||
QY 361 cgggttcgagttgaggaagcggcggaagccaagtaaaagaagaagaatgaaagaagagc 420
|||||
DB 829 CGGTTGAGAGTTGGGAGACGGCCGGAAGCCAGTMAAAGAGAGAAATGGAAGAAAGGAGC 888
|||||
QY 421 ctgctcttcagactacagagacatcagcagcagcaggaagaaagacacgctcgccgagtc 480
|||||
DB 889 CTGCTGTCCGACTTACAGACATCACGACGAGGAGAGGAGACCGACTCGCGCGGAGTCC 948
|||||
QY 481 tgcgcagcagaatcaggtctcaggttctgaggaagaaagaggaaggaaggaagagag 540
|||||
DB 949 TCGTTCAGCAAGATCAGGCTCAGGTTCTGAGGAAGAAAGAGAGAGAGAGAGAGAG 1008
|||||
QY 541 gagggaaggaagcagcagtgaaagaatcagagaggaaggaaggaaggaaggaaggaagc 600
|||||
DB 1009 GAGGAAGGAGACCCAGTGAAGATCAGAGAGAGAAAGAGAGAGAGAGAGAGAGAGAC 1068
|||||
QY 601 ggcagcaactctgaaggaagcagcagcagcttcgcgaagaagtaagtgaagaagaatg 660
|||||
DB 1069 GGCAGCAACTCTGAGAGAGGAGCATCAGAGCAGTGTCCGGAAGAAGTAAGTACGAGAAAGATG 1128
|||||

RESULT 5
AF067525 2329 bp mRNA PRI 07-NOV-1998
LOCUS Homo sapiens PITSLRE protein kinase beta SV13 isoform (CDC2L2)
DEFINITION
AF067525 mRNA, complete cds.
ACCESSION
AF067525.1 GI:3850329
VERSION
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2329)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,

TITLE Ambros,P.F. and Kidd,V.J.
JOURNAL Duplication of a genomic region containing the Cdc2L1-2 and
MEDLINE MRP21-22 genes on human chromosome 1p36.3 and their linkage to D122
REFERENCE Genome Res. 8 (9), 929-939 (1998)
AUTHORS 2 (bases 1 to 2329)
Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
RESEARCH Location/Qualifiers
FEATURES
source
1..2329
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2329
/gene="CDC2L2"
68..1654
/note="corresponding genomic sequence deposited as GenBank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PITSLRE protein kinase beta SV13 isoform"
/protein_id="AAC72090.1"
/db_xref="GI:3850330"

BASE COUNT 634 a 605 c 738 g 352 t
ORIGIN

Query Match 99.5%; Score 656.8; DB 9; Length 2329;
Best Local Similarity 99.7%; Pred. No. 4.9e-112;
Matches 658; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 caggaagcgtcgaaagacacatcgagaagacaagataaagctcgccgggaatgggaaga 60
|||||
DB 251 CACGAAGCTCGAAGACGACATCGAAGAAAGACAGATTAACCTCGCCGGGATGGAAAGA 310
|||||
QY 61 cagaagaagaaagaaatgcaaggaagcattccagagagaagaaggggaatgagcgatg 120
|||||
DB 311 CAGAAGAGAAAGGAAATGCGCAAGGAGAGCATTCAGAGAGAAAGAGGGAATGATGCGT 370
|||||
QY 121 tgcctcttcagagaccgcttgagacagtttagaagaagcggagcggagcggcaaatg 180
|||||
DB 371 TGCCTCTTCAAGGACCGCTTGAGACAGTTGAAAGAAAGCGGAGCGGAGCCGAAGATG 430
|||||
QY 181 cgggaagcagcaggaaggaagcagcgggaagcaggaaggaagcggcgagcgggaagcg 240
|||||
DB 431 CGGAGCAGCAGAGAGAGAGCAGCGGAGCAGAGAGAGCGGAGCGGCGGAGAGAGAGCG 490
|||||
QY 241 cggcaagaagcggagagcccgcaaggaagtgctcgcacatccagcaagcagtgagagagc 300
|||||
DB 491 CGCAAGAGCGGAGGCGCCGAGGAGGAAGTGTGACATCACGGAACGATGAGAGAGAGAC 550
|||||
QY 301 taacagcagcaaaagtgaagaagccaagcagctgagtgagcagccgctcgccgagcgag 360
|||||
DB 551 TACAGCGACAAAGTGAAGAGCAGCCACTGGAAGTGCAGCCCGCTCGCGCGGAGAG 610
|||||
QY 361 cgggttcgagttgaggaagcggcggaagccaagtaaaagaagaagaatgaaagaagagc 420
|||||
DB 611 CGGTTGAGAGTTGGGAGAGCGCGGAGAGCCAGTAAAGAAAGAAATGGAAGAAAGGAGC 670
|||||

QY	361	cggctcgtagctctgggaagacggcgcgaagccacgcaaaaagaaatctgaagaagagac	420
	823	CGGTTCCAGTTGGGAGACGGCCGGAAGCCAGTAAAGAGAAATAATGAAAGAAAGGAC	882
QY	421	ctgcctgcacactctacagacatccagcagcagagagaaagccagctctggccgagctc	480
Db	883	CTGCTGTCCGACTTACAGGACATCACCGACCGAGAGAAACCACTCGCCGCGATCC	942
QY	481	tcgtcagcagaatccaggtctcaggtctctctggagaagagggagaaagaaagaaagag	540
Db	943	TCGTCACCAAAATCTAGGCTCTAGGTTCTTGAGAAAGAGAGAGAGAGAGAGAGAG	1002
QY	541	gaggaagggagacacacagctgaaagatctagaagaggaagaaagaaagagagagagacc	600
Db	1003	GAGGAAGGAGACACACAGTGAAGAAATCAGAGGAGAAAGAGAAAGAGAGAGAGAGACC	1062
QY	601	ggcagcaactctgaagagacatcaagagcagctctccgaaagaagtgaatgagaaagaatg	660
Db	1063	GGCGACAACTCTGAGGAGGCGCATGAGCAGCTCGCCGAAGAAAGTAAGTGAGGAACAAATG	1122

[illegible]

REFERENCE 1 (bases 1 to 2471)
AUTHORS Beyaert, R., and Cornelis, S.
TITLE Internal ribosome entry site (IRES), vector containing same and uses thereof
JOURNAL Patent: WO 004896-A 3 03-AUG-2000;

JOURNAL
FALCINI: WO 0044830 A 5 03 NOV 2000,
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)

FEATURES	source	location/Qualifiers
		1..2471
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
misc-feature		1..2471
		/note="The IRS-activity containing sequence PTSLRE protein kinase (p110ptslre) (isoform alfa 2-2)"
BASE COUNT	697 a	601 c 809 g 364 t
ORIGIN		

Query Match	100.0%	Score 660:	DB 6:	Length 2471:
Best Local Similarity	100.0%	Pred. NO. 1.3e-112:		
Matches 660: Conservative	0:	Mismatches	0:	Gaps 0

QY	1	cagaagctcggaacgcacatcsgaagaacagatataagctctgcggaatcgggaaga	60
Db	469	CACGACCTCGGAAACGACATCTGAGGAACAGGATTAACCTGCGCGGATGGANA	528
QY	61	cagaagagaaggaatctgcgaagagacatcccgagagaaaggggaatga.tgctgtg	120
Db	529	CAGAGAGACAGGAATGTGGCAGAGGACATCCAGGAGCAAAAGGGGATGTGCGCTG	588
QY	121	tgctcttcgaaggacgcgtcttgagcagtttgaaaaagaaagcggagcgggaagcgaatgtg	180
Db	589	TGCCCTTCACGGAGACCGCTTGACACTTAAGAAAGGAAGCGGAGCGGAGCCCAAGATG	648
QY	181	cggagacagaaagagcagcggagagcaagaagagacgcagcggcggcggagagcgg	240
Db	649	CGGAGCAGCAAGACAGACACGCGGAGCAGAGGAGCGCGAGCGCGCGGAGGAGCGG	708
QY	241	cgcgaagagcggagagcccgacagggaaagtctctcacaataccgaaacgatgaaagagac	300
Db	709	CCCAAGAGACCGGAGCCCGCAGGGAATTGTCTCACTACATCCCAACGATGAGAGAGAC	768

QY	301	taagagcaaaatgaaagccagccacactgtgagtcgacgcccgtctygcgcgcgcggag	360
Db	769	TACAGCACAAGTGAAGACCACACTGAGTGTGCAAGCCCGCTGGCCGCCGCGGAG	828
QY	361	cggttcagatgtggaagacgycgcggaaagccagtaaaagaagaaatgaaagaagagac	420
Db	829	CGGTTCCAGTTGGAGAGACGGCCGGAAAGCCAGTAAAAAGAGACAAATTCAGACAAAGCGAC	888
QY	421	ctgctgttcgacttaacagagacatcagagaaagaagaaagaagacagcctcgcgcgaatcc	480
Db	889	CTGCTGTCCAGTTCACAGGACATCAGGAGACGAGAGGAGACACAGCTCGGCCAGTCC	948
QY	481	tcgtcagcagaatcaagcctcaagttctctgaaagaagaagaggaaggaagaaagaagag	540
Db	949	TCTGTACAGAAATCAGGCTCAGGTTTGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAG	1008
QY	541	gaagaaagagacacccaatgaaagaatcaagaagaagaagaagaagaagaagaagaagacc	600
Db	1009	GAGGAAGGACCCACAGTGAAGATTCAGAGAGAGAAAGAGAAAGAGAGAGAGAGAGACC	1068

[illegible]

RESULT	4
LOCUS	HSU04816
DEFINITION	Human protein kinase PTKSLK alpha 2-2 mRNA, complete cds.
ACCESSION	U04816.1
VERSION	GI:507159
	08-JUL-1994

SOURCE

ORGANISM	REFERENCE
<i>Homo sapiens</i>	1 (bases 1 to 2471)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
Xiang, J., Lahti, J. M., Grenet, J. A., Easton, J. B. and Kidd, V. J.	
Molecular cloning and expression of alternatively spliced PITSARE	

JOURNAL J. BIOL. Chem. 269, 15786-15794 (1994)
MEDLINE 94253170
REFERENCE 2 (bases 1 to 2471)
AUTHORS Kidd, V.J.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's

FEATURES

```

source
1. .2471
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1p36-2"
/cell_line="Hela S3"
/tissue_type="epitheloid"
/tissue_type="cervix"
/clone_lib="Hela cDNA library"
/dev_stage="adult"
1. .2471
mrna
1. .2471
112. .2445
cds

```

```

/codon_start=1
/product="PITSLRE alpha 2-2"
/protein_id="AA19582.1"
/db_xref="GI:507160"
/translation="WGDEKDSWKVKTLDLLEIQEKKRKEDEKAEIKRLKNSDRDSK
RDSLEGEIRDHCMETITRNSPYRREDSMERGEDSLAKPQOQSRKRVVHRKD
EKRKKEHARVYKKEHEHRRKRIHREEDMRKREGEERKREMAHEHSRRKRDGVO
LFRDLDEKRRERERKREDEQKQDQKQKRRERKAEERKREARREVSNAHTMEE
DYSKQVKAHSRSPRPKRPREFGLDQKRVKKEKMRPLDLIDIOSERTKST
AASSAEGSGSGSEEEEEEEEEEEBEGSTSESESEEEEEEEBETSNSEASQAEVY
SEEMSEDERENNENHLLVPEPSRFDQDGSSEEEAEVGGTQSAVLTEGDIYPPSS
PALSTELKQELPKYLPALQDQSRVEEPCQCNRIIEGTVYGVYAKKAKTDEIVALKR

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:54:37 ; Search time 2399.05 Seconds
(without alignments)
4538.521 Million cell updates/sec

Title: US-09-915-060-6
Perfect score: 660
Sequence: 1 cagcaacgtcggaaacgaca.....aagtaagtgaagaagaatg 660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: GenBank:*
- 2: gb_ba:*
- 3: gb_lin:*
- 4: gb_ov:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htgo_hum:*
- 31: em_htgo_inv:*
- 32: em_htgo_rod:*
- 33: em_htg_hum:*
- 34: em_htg_inv:*
- 35: em_htg_rod:*
- 36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	660	100.0	660	6	AX033425	AX033425 Sequence
2	660	100.0	2465	6	HSU07705	U07705 Human prote
3	660	100.0	2471	6	AX033422	AX033422 Sequence
4	660	100.0	2471	9	HSU04816	U04816 Human prote
5	656.8	99.5	2329	9	AF067525	AF067525 Homo sapi
6	656.8	99.5	2436	9	AF067520	AF067520 Homo sapi
7	596	90.3	2323	9	AF067518	AF067518 Homo sapi
8	596	90.3	2409	9	AF067521	AF067521 Homo sapi
9	596	90.3	2439	9	AF067519	AF067519 Homo sapi
10	596	90.3	2444	9	HSU04817	U04817 Human prote
11	596	90.3	2471	9	HSU07704	U07704 Human prote
12	596	90.3	2477	9	HSU04824	U04824 Human prote
13	577	87.4	2486	9	AF067512	AF067512 Homo sapi
14	577	87.4	2500	9	AF067514	AF067514 Homo sapi
15	577	87.4	2533	9	AF067516	AF067516 Homo sapi
16	547	82.9	2362	9	AF067523	AF067523 Homo sapi
17	547	82.9	2448	9	AF067522	AF067522 Homo sapi
18	528	80.0	2525	9	AF067515	AF067515 Homo sapi
19	520	78.8	2544	9	AK000081	AK000081 Homo sapi
20	468	70.9	468	6	AX033424	AX033424 Sequence
21	397.8	60.3	3161	10	MUSCDPK	L37092 Mus musculu
22	395.4	59.9	2340	9	HSU04818	U04818 Human prote
23	376.4	57.0	2349	9	AF067517	AF067517 Homo sapi
24	371.4	56.3	1715	9	AF174497	AF174497 Homo sapi
25	252.8	38.3	1936	10	MUSP58GTA	M58633 Mouse p58/G
26	252.8	38.3	2226	9	AF067513	AF067513 Homo sapi
27	222	33.6	222	6	AX033420	AX033420 Sequence
28	222	33.6	222	6	AX033421	AX033421 Sequence
29	166	25.2	110608	9	HS283E3	AX031282 Human DNA
30	165	25.0	297	9	CDCC2LIS06	AF092429 Homo sapi
31	161.8	24.5	297	9	CDCC2LIS06	AF092427 Homo sapi
32	155.4	23.5	592	9	CDCC2LIS08	AF080689 Homo sapi
33	138.4	21.0	712	9	AF067524	AF067524 Homo sapi
34	134.4	20.4	480	9	CDCC2LIS08	AF080678 Homo sapi
35	127.4	19.3	532	9	CDCC2LIS07	AF092428 Homo sapi
36	127.4	19.3	533	9	CDCC2LIS07	AF092430 Homo sapi
37	105.4	16.0	539	9	CDCC2LIS05	AF080688 Homo sapi
38	105.4	16.0	761	9	CDCC2LIS05	AF080697 Homo sapi
39	100	15.2	551	9	AF067527	AF067527 Homo sapi
40	100	15.2	700	9	AF067529	AF067529 Homo sapi
41	99.4	15.1	7218	6	166494	166494 Sequence 14
42	99.2	15.0	318930	2	AC073495	AC073495 Mus muscu
43	98	14.8	579	9	AF067526	AF067526 Homo sapi
44	98	14.8	728	9	AF067528	AF067528 Homo sapi
45	96.4	14.6	179149	9	AC009785	AC009785 Homo sapi

ALIGNMENTS

RESULT 1
AX033425 AX033425 660 bp DNA PAT 21-SEP-2000
LOCUS Sequence 6 from Patent WO0044896.
DEFINITION AX033425
ACCESSION AX033425
VERSION AX033425.1 GI:10280186
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Beyaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
JOURNAL Patent: WO 0044896-A 6 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)
FEATURES
source Location/Qualifiers
1..660


```

RESULT 14
LOCUS   CNS07C7R 1117 bp DNA GSS 08-JUL-2001
DEFINITION T3 end of clone XBC0A002F11 of library XBC0A from strain CBS 767
of Debaryomyces hanseni, genomic survey sequence.
ACCESSION AL438829
VERSION   AL438829
KEYWORDS  AL438829.1 GI:12222242
SOURCE    GSS.
ORGANISM  Debaryomyces hanseni.
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS   Lepingle, A., Casaregola, S., Neuvéglise, C., Bon, E., Nguyen, H.,
          Artiguenave, F., Wincker, P., and Galliardin, C.
          Genomic exploration of the hemiascomycetous yeasts: 14.
          Debaryomyces hanseni var. hanseni
          FEMS Lett. 487 (1), 82-86 (2000)
          20584724
TITLE      2 (bases 1 to 1117)
JOURNAL    Soucier, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
MEDLINE     Boletín-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,
AUTHORS     de-Montigny, D., Dujon, B., Durrens, P., Lepingle, A., Liorente, B.,
          Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
          Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,
          Wincker, P. and Weissenbach, J.
          Genomic exploration of the hemiascomycetous yeasts: 1. A set of
          yeast species for molecular evolution studies
          FEMS Lett. 487 (1), 3-12 (2000)
          20584711
TITLE      3 (bases 1 to 1117)
JOURNAL    Genoscope.
MEDLINE     Direct Submission
AUTHORS     Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,
          2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :
          sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)
          This GSS is part of a random genomic sequencing program of thirteen
          yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
          exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
          Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
          lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
          anomala, Debaryomyces hanseni var. hanseni, Pichia sorbitophila,
          Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
          5 kb were prepared and both extremities were sequenced. See
          keywords for description of this sequence and for the sequence of
          the other extremity of this insert.
FEATURES
  source
    1. 1117
      /organism="Debaryomyces hanseni"
      /strain="CBS 767"
      /variety="hanseni"
      /db_xref="taxon:4959"
      /clone="XBC0A002F11"
      /clone_1ib="XBC0A"
      /note="end : T3"
      /note=">580"
      /note="Similar to Saccharomyces cerevisiae ORF YDR084c [
      similarity to hypothetical C.elegans protein ]"
      /evidence=not_experimental
BASE COUNT 351 a 159 c 195 g 409 t 3 others
ORIGIN
Query Match 54.2%; Score 20.6; DB 13; Length 1117;
Best Local Similarity 85.2%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 7 agagaattcgaagtgaactctt 33
    1111 1111 11111111 1111
DB 135 AGAGCAATTCGATGTCAGCATCTTT 109

```

```

RESULT 15
LOCUS   AQ402410 546 bp DNA GSS 13-MAR-1999
DEFINITION HS_5060_A1_F10_T7A RPT-11 Human Male BAC Library Homo sapiens
genomic clone Plate=636 Col=19 Row=K, DNA sequence.
ACCESSION AQ402410
VERSION   AQ402410.1 GI:4413322
KEYWORDS  GSS.
SOURCE    human.
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 546)
          Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
          Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
          Hood, L.
          Sequence-tagged connectors: A sequence approach to mapping and
          scanning the human genome
          Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
          99380589
TITLE      High Throughput Sequencing Center
JOURNAL    Contact: Mahairas GG, Wallace JC, Hood L
MEDLINE     University of Washington
AUTHORS     401 Queen Anne Avenue North, Seattle, WA 98109, USA
          Tel: (206) 616-3618
          Fax: (206) 616-3887
          Email: jwallace@u.washington.edu
          Clones are derived from the human BAC library RPT-11. For BAC
          library availability, please contact Plier de Jong
          (plier@dejong.med.buffalo.edu). Clones may be purchased from
          BACPAC Resources (http://bacpac.med.buffalo.edu/ordering-bac.htm)
          or from Resear h Genetics (info@resgen.com). BAC end Web Server:
          http://www.htsc.washington.edu
          Plate: 636 row: K column: 19
          Seq primer: T7
          Class: BAC ends
          High quality sequence stop: 546.
FEATURES
  source
    1. 546
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="Plate=636 Col=19 Row=K"
      /clone_1ib="RPT-11 Human Male BAC Library"
      /sex="male"
      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
      Male blood DNA was isolated from one randomly chosen donor
      and partially digested with a combination of EcoRI and
      EcoRI Methylase. Size selected DNA was cloned into the
      pBACe3.6 vector at EcoRI sites"
BASE COUNT 191 a 95 c 91 g 165 t 4 others
ORIGIN
Query Match 53.7%; Score 20.4; DB 13; Length 546;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 ctctagaggaattcgaagtgaactctt 32
    1111111111111111 11
DB 45 CTTCTAGAGGAATTCGACGATGCAATCAATT 16

```

Search completed: December 4, 2001, 18:14:29
 Job time: 6770 sec

REFERENCE	Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
AUTHORS	1 (bases 1 to 990)
TITLE	Malpertuy, A., Llorente, B., Blandin, G., Artiguenave, F., Wincker, P. and Dujon, B.
JOURNAL	Genomic exploration of the hemiascomycetous yeasts: 10.
MEDLINE	Kluyveromyces thermotolerans
REFERENCE	FEBS Lett. 487 (1), 61-65 (2000)
AUTHORS	20584720
TITLE	2 (bases 1 to 990)
JOURNAL	Souciat, J. L., Aigle, M., Artiguenave, F., Blandin, G.,
MEDLINE	Botolin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,
REFERENCE	de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,
AUTHORS	Malpertuy, A., Neuvéglise, C., Ozler-Kalogeropoulos, O., Potier, S.,
TITLE	Saurin, M., Tekata, F., Toffano-Nioche, C., Wesolowski-Douvet, M.,
JOURNAL	Wincker, P. and Weissenbach, J.
MEDLINE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of
REFERENCE	yeast species for molecular evolution studies
AUTHORS	FEBS Lett. 487 (1), 3-12 (2000)
TITLE	3 (bases 1 to 990)
JOURNAL	Genoscope.
MEDLINE	Direct Submission
REFERENCE	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
AUTHORS	2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
TITLE	bept@genoscope.cns.fr - Web : www.genoscope.cns.fr)
JOURNAL	This GSS is part of a random genomic sequencing program of thirteen
MEDLINE	yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
REFERENCE	exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
AUTHORS	Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
TITLE	lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
JOURNAL	angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
MEDLINE	Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
REFERENCE	5 kb were prepared and both extremities were sequenced. See
AUTHORS	keywords for description of this sequence and for the sequence of
TITLE	the other extremity of this insert.
JOURNAL	Location/Qualifiers
MEDLINE	1. 990
REFERENCE	/organism="Kluyveromyces thermotolerans"
AUTHORS	/strain="CBS 6340"
TITLE	/db_xref="taxon:4916"
JOURNAL	/clone="AY0AA007F09"
MEDLINE	/clone_lib="AY0AA"
REFERENCE	/note="end : T7"
AUTHORS	<241. >471
TITLE	/note="similar to P78771 [Schizosaccharomyces pombe,
JOURNAL	unknown protein]"
MEDLINE	/evidence="not-experimental"
REFERENCE	BASE COUNT 219 a 272 c 265 g 231 t 3 others
AUTHORS	ORIGIN
TITLE	misc_feature
JOURNAL	<241. >471
MEDLINE	/note="similar to P78771 [Schizosaccharomyces pombe,
REFERENCE	unknown protein]"
AUTHORS	/evidence="not-experimental"
TITLE	BASE COUNT 219 a 272 c 265 g 231 t 3 others
JOURNAL	ORIGIN
MEDLINE	Query Match 55.3%; Score 21; DB 13; Length 990;
REFERENCE	Best Local Similarity 73.0%; Pred. No. 2.3e+02;
AUTHORS	Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
TITLE	Oy 1 tgcctagagagatcgaagtacagataactcttgagc 37
JOURNAL	
MEDLINE	Db 176 ttcgccacagcgcgctctgaagtagcagctactcttgccg 140
REFERENCE	RESULT 10
AUTHORS	BG619393/C 734 bp mRNA EST 18-APR-2001
TITLE	LOCUS 602619138P1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4732929 5',
JOURNAL	DEFINITION mRNA sequence.
MEDLINE	BG619393
REFERENCE	ACCESSION BG619393.1 GI:13670764
AUTHORS	VERSION EST.
TITLE	KEYWORDS human.
JOURNAL	SOURCE Homo sapiens
MEDLINE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	REFERENCE 1 (bases 1 to 734)

AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: rgapubs@email.nih.gov Tissue Procurement: CLONTECH Laboratories, Inc. cDNA Library Preparation: CLONTECH Laboratories, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1CMI593 row: k column: 10 High quality sequence stop: 701. Location/Qualifiers			
FEATURES	1. .734			
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4732929" /clone_id="NIH_MGC_79" /lab_host="DH10B (T1 phage-resistant)" /note="Organ: placenta; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgcctcgcc); Site_2: SfiI (ggcctatggcc) : 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGGCCGACGCGCCGACGATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, Ca). Note: this is a NIH_MGC Library."			
BASE COUNT	223 a	139 c	127 g	245 t
ORIGIN				
Query Match	54.7%; Score 20.8; DB 11; Length 734;			
Best Local Similarity	78.1%; Pred. No. 2.6e+02;			
Matches	25;	Conservative	0;	Mismatches 7; Indels 0; Gaps 0;
Oy	1	tgctctagagaatcgaagtgcacgatctt 32		
Db	477	TGCTCTAAAAAATTCGAAGTMAACGTATTT 446		
RESULT 11				
LOCUS	AC0696864	430 bp	DNA	GSS
DEFINITION	HS-5526_A2_D03-SF6E RPl-11 Human Male BAC Library Homo sapiens genomic clone Plate-1102 Col-6 Row-G, DNA sequence.			
ACCESSION	AC0696864			
VERSION	AC0696864.1	GI:5387112		
KEYWORDS	GSS.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 430) Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Ketter,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.			
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome			
JOURNAL	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)			
MEDLINE	99380589			
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones are derived from the human BAC library RPl-11. For BAC			

`/db_xref="taxon:10090"`
`/clone="UGC1M0569E11"`
`/clone_lib="Mouse 10kb plasmid UGC1M library"`
`/sex="Male"`
`/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"`
`/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (911473211419b) (AF123072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."`

BASE COUNT
ORIGIN

147 a	123 c	122 g	106 f
-------	-------	-------	-------

Query Match	55.3%;	Score 21;	DB 13;	Length 498;
Best Local Similarity	73.0%;	Pred. No. 2e+02;		
Matches 27; Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

```

QY      1  tgcctcagaggaattcgaaagtgcagataactttggcg  37
          |||||  |  |||  |||  |||  |||||
Db      327  TGCTCAAAAGCAATCCACGATAGCGATGGTCTTGGCG  291

```

RESULT 7					
AU094789/c					
LOCUS	AU094789	618 bp	mRNA	EST	30-JUN-2000
DEFINITION	AU094789 Rice panicle shorter than 3cm Oryza sativa cDNA clone E31652, mRNA sequence.				

ACCESSION	AU094/89	GI:8857471
VERSION	AU094789.1	
KEYWORDS	EST.	
SOURCE	Oryza sativa.	
ORGANISM	Oryza sativa	

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE	1 (bases 1 to 618)
AUTHORS	Sasaki,T. and Yamamoto,K.
TITLE	Rice cDNA from panicle (2000)
JOURNAL	Unpublished (2000)
COMMENT	Contact: Takuji Sasaki

National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasakia@affrc.go.jp, URL: <http://rpg.dna.affrc.go.jp/>
PROJECT = 'RGP'.
E31662-62.

FEATURES	Location/Qualifiers
source	1. .618

```

/organism="Oryza sativa"
/strain="Nippobaire"
/db_xref="taxon:4530"
/clone="E31662"
/clone_id="Rice panicle shorter than 3cm
/dev_stage="shorter than 3cm"
/note="Organ: panicle"

```


ORIGIN

Query Match	55.3%;	Score 21;	DB 10;	Length 618;
Best Local Similarity	73.0%;	Pred. No. 2.1e+02;		
Matches 27;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;

```

QY      2  gctctagaggaattctgaagtgcgatacttttgccg 38
          | | | | | | | | | | | | | | | | | |
Db      60  GGTGTGACGAGGACGAAGTCCGATGANTTCGGCGC 24

```

RESULT	8
AL574604/c	
LOCUS	763 bp mRNA
DEFINITION	AL574604 LTI_NF1006_P12 Homo sapiens cDNA clone CS0D1056YJ18 3'

ACCESSION	AL574604
VERSION	AL574604.1
	GI:129349666

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
1 (bases 1 to 763)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.

AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope

Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

FEATURES

SOURCE

Location/Qualifiers
1. .763

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDD1056YJ18"
/clone_lib="TRL_NFL006_PL2"
/tissue_type="placenta"
/notice="vector: pcwvSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pcwvSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : liang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

BASE COUNT 238 a 119 c 164 g 237 t 5 others

ORIGIN

Query Match	55.3%;	Score 21;	DB 10;	Length 763;
Best Local Similarity	77.4%;	Pred. No. 2.2e+02;		
Matches	24;	Conservative	1;	Mismatches 6;
			Indels	0;
			Gaps	0;

Qy	4	tctagaggaattcgaagtgcagataactttg	34
		:	
Db	114	tcttgacgaattmgtagagagataacttttg	84

RESULT 9

CNS06YA2/C	
LOCUS	CNS06YA2 990 bp DNA GSS 06-JUL-2001
DEFINITION	T7 end of clone AY0A007F09 of library AY0A from strain CBS 6340 of Kluyveromyces thermotolerans, genomic survey sequence.

ACCESSION	AL420/68
VERSION	AL420768.1
	GI:12203955

KEYWORDS GSS.
SOURCE *Kluyveromyces thermotolerans*.
ORGANISM *Kluyveromyces thermotolerans*
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes

KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 906)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ANCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigrl.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McQuinn, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgenc, C., Vogt, J.L., Walker, M.A., Zhang, L., H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 26 Row: d Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 3850303
This clone has the following problem: incomplete processing.
Location/Qualifiers
1. .906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:412154"
/tissue_type="Muscle, Rhabdomyosarcoma"
/clone_id="NHL_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

BASE COUNT 317 a 170 c 307 g 112 t
ORIGIN

Query Match 57.9%; Score 22; DB 12; Length 906;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 gaagtgaacatactttggcgc 38
|||||
Db 12 GAAGTGACATATCTTGGCC 33

RESULT 5
LOCUS A1728511 564 bp mRNA EST 11-JUN-1999
DEFINITION BNGH110926 Six-day cotton fiber Gossypium hirsutum cDNA 5' similar to FERROCHELATASE PRECURSOR (PROHEME FERRO-LYASE) (HME SYNTHETASE) g114749681gnl|PID|d1005644 (D26106) ferrochelatase [Gossypium sativus], mRNA sequence.
ACCESSION A1728511
VERSION A1728511.1 GI:5047363
KEYWORDS EST.
SOURCE upland cotton.
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 564)
AUTHORS Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.
TITLE ESTs from developing cotton fiber
JOURNAL Unpublished (1999)
COMMENT Contact: Ben Burr
Brooklyn Department
Brooklyn National Laboratory
Upton, NY 11973, USA
Tel: 516-344-3396
Fax: 516-344-3407
Email: burreb@nslu1.bnl.gov
Seq primer: T3 Primer.

FEATURES
source location/Qualifiers
1. .564
/organism="Gossypium hirsutum"
/cultivar="Acala Maxxa"
/db_xref="taxon:3635"
/clone_lib="Six-day Cotton fiber"
/tissue_type="Immature fiber"
/dev_stage="Six days post anthesis"
/lab_host="XLI-Blue"
/note="Vector: pBluescript II KS+"

BASE COUNT 152 a 107 c 144 g 160 t 1 others
ORIGIN

Query Match 56.3%; Score 21.4; DB 10; Length 564;
Best Local Similarity 80.6%; Pred. No. 1.4e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 4 tctagaagattcgaatgacatactttg 34
|||||
Db 273 TCTAGAGATTGGAATGATACCTTG 303

RESULT 6
LOCUS A2768917 498 bp DNA GSS 16-FEB-2001
DEFINITION 1M0569E11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0569E11 F, DNA sequence.
ACCESSION A2768917
VERSION A2768917.1 GI:12888508
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 498)
AUTHORS Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausen, A. and Wright, D., Weils, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10Kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0569 row: E column: 11
Seq primer: CGTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 498.
Location/Qualifiers
1. .498
/organism="Mus musculus"
/strain="C57BL/6J"

BASE COUNT	ORIGIN
82 a	105 c 134 g 88 t 2 others

		58.4%;	Score 22.2;	DB	10;	Length	411;
Query Match		77.1%;	Pred.	No. 66;			
Best Local Similarity		0;	Mismatches	8;	Indels	0;	Gaps
Matches	27;	Conservative					
Qy	3	ctctagagaattcgaatgacagataactttggcg	37				
Db	308	CTGTAGAAGGTTTCATCTTGACGACTTGGCGC	274				

RESULT	2
LOCUS	AA213378
DEFINITION	AA213378 488 bp mRNA EST 12-MAR-1998
ACCSSION	Z693010.s1 Stratagene hNT neuron (#937253) Homo sapiens cDNA clone
VERSION	IMAGE:649554 3', mRNA sequence.
KEYWORDS	AA213378.1 GI:1812096
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 488)
AUTHORS	Hillier,L., Allen,M., Bowles,L., Dubouque,T., Geisels,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lemon,G., Maira,M., Martin, J., Moore,B., Schellenberg,K., Sepcoe,M., Tan,F., Theising,B., White,I., Wyllie,T., Waterston,R. and Wilson,R.
TITLE	WashU-NCI Human EST Project
JOURNAL	unpublished (1997)
COMMENT	Contact: Wilson RK

FEATURES

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewartson.wustl.edu

This clone is available royalty-free through LINL ; contact the
IMAG Consortium (info@image.linn.gov) for further information.
Insert Length: 1974 Std Error: 0.00
Seq. primer: -41m3 fwd. ET from Amersham
High quality sequence stop: 281.
Location/Qualifiers

```

source
1. .488
/organism="Homo sapiens"
/db_xref="GDB:5278834"
/db_xref="taxon:9606"
/clone="IMAGE:649554"
/clone_1id="Stratagene hNT neuron (#937233)"
/dev_stage="hNT neurons"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt.
Size: 1.5 kb; Uni-ZAP XR Vector: 5' adaptor sequence: 5
GAATTCGACGACGAG 3' -3' adaptor sequence: 5
CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT      84 a      135 c      169 g      93 t      7 others
ORIGIN

```

Query Match	57.9%;	Score 22;	DB 10;	Length 488;
Best Local Similarity	100.0%;	Pred. No. 81;		
Matches 22;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	17	gaagtgacataactttggcgc	38	
Db	6	gaagtgacacatcrrttggcgc	27	

RESULT	3				
LOCUS	BG827010				
DEFINITION	BG827010	640 bp	mRNA	EST	22-MAY-2001
	602749042n1	NH_MGC_17	Homo sapiens	cdna clone	IMAGE:4901723 5',
ACCESSION	BG827010				
VERSION	BG827010.1	GI:14174597			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 640)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cs9apbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: L10CMT197 row: d column: 12
High quality sequence stop: 600.

FEATURES	SOURCE
Location/Qualifiers	1. .640
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="IMAGE:4901723"	
/clone_lib="NH.MGC.17"	
/tissue_type="rhabdomyosarcoma"	
/lab_host="DH10B (phage-resistant)"	
/note="Organ: muscle; Vector: pORF7; Site_1: EcoRI; Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Size-selected >500bp for average insert size 1.8kb library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using Zap-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."	
BASE COUNT	218 a 125 c 199 g 97 t 1 others
ORIGIN	

Query Match	57.9%	Score 22:	DB 11:	Length 640:
Best Local Similarity	100.0%	Pred. No. 84:		
Matches 22:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
OY	17	gaagtgaacatactttggcgc	38	
Db	4	GAAGTGACGATACCTTTGGCGC	25	
RESULT	4			
LOCUS	BC009375			
DEFINITION	BC009375	906 bp	mrna	HTC
				09-JUL-2001
				Homo sapiens, similar to cell division cycle 2-like 1 (PTSLREB
				proteins), clone IMAGE:4121554, mRNA.
ACCESSION	BC009375			
VERSION	BC009375.1	GI:14627288		

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:14:26 ; Search time 3881.49 Seconds
(without alignments)
105.202 Million cell updates/sec

Title: US-09-915-060-7
Perfect score: 38
Sequence: 1 tgcctcagaggaattcgaagtgcacgactttgagcgc 38

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estopl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_est1:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rpd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22.2	58.4	411	10	AA159840 z056907.r
2	22.2	57.9	488	10	AA213378 z056907.s
3	22.2	57.9	640	11	BG827010 602749042
4	22.2	57.9	906	12	BC009375 Homo sapi
5	21.4	56.3	564	10	AI728511 BNGH1109
6	21.4	55.3	498	13	AZ768917 IM0569E11
7	21.4	55.3	618	10	AU094789 AU094789
8	21.4	55.3	763	10	AL574604 AL574604
9	21.4	55.3	990	13	CNS06YA2 AL420768
10	20.8	54.7	734	11	BG619393 602619138
11	20.6	54.2	430	13	A0696864 HS_5526_A
12	20.6	54.1	13	13	A0511435 HS_5214_A

c 13	20.6	54.2	642	10	AI638929 AEMTAL50
c 14	20.6	54.2	1117	13	CNS07C7R AL438829
c 15	20.4	53.7	546	13	AA0402410 HS_5060_A
c 16	20.2	53.2	300	11	C07641 C07641
c 17	20.2	53.2	352	11	R03855 PK12609.s1
c 18	20.2	53.2	660	13	AZ471544 AZ471544
c 19	20.2	53.2	666	13	AO545121 CITR1-E1
c 20	20.2	53.2	822	10	AM9833246 HVSME900
c 21	20.2	53.2	881	11	BF244937 601864340
c 22	20.2	52.6	306	10	AI524605 t043h04.x
c 23	20.2	52.6	327	10	AA663181 ab74c01.s
c 24	20.2	52.6	432	11	BF566267 UI-R-BT1.s
c 25	20.2	52.6	491	10	BE511204 946058F05
c 26	20.2	52.6	510	11	BG411126 EM1_26_A0
c 27	20.2	52.6	516	13	AO721424 HS_5559_A
c 28	20.2	52.6	613	10	AA455707 707090809
c 29	20.2	52.6	684	13	CNS01STS AL165673
c 30	20.2	52.6	798	11	BG321018 zmo4_02h0
c 31	20.2	52.6	892	11	BG321059 zmo4_02h0
c 32	20.2	52.6	1403	11	BG114146 602285564
c 33	19.8	52.1	140	10	AA840946 MB3D6A1F
c 34	19.8	52.1	414	13	A0145635 HS_2218_B
c 35	19.8	52.1	525	13	TA293601P AL485274
c 36	19.8	52.1	527	11	BG447599 NF117B10S
c 37	19.8	52.1	570	10	AI438619 486012P06
c 38	19.8	52.1	616	10	AM687972 NF001P06S
c 39	19.8	52.1	636	10	AM689592 NF022B03S
c 40	19.8	52.1	640	10	AI438623 486012S01
c 41	19.8	52.1	654	10	AM695839 NF099B08S
c 42	19.8	52.1	664	10	BE204210 EST396886
c 43	19.8	52.1	719	11	BF641446 NF056D11I
c 44	19.8	52.1	719	11	BG645842 EST507461
c 45	19.6	51.6	300	11	C59696 C59696 yuji

ALIGNMENTS

RESULT 1
LOCUS AA159840/c
DEFINITION z056907.r1 Stratiagene pancreas (#937208) Homo sapiens cDNA clone IMAGE:590940 5' similar to SW:DBRL_YEAST P24309 LARIAT DEBRANCHING ENZYME ; mRNA sequence.

ACCESSION AA159840
VERSION AA159840.1 GI:1734985

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 411)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Madis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Warr, M.

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 330.

TITLE JOURNAL MEDLINE

COMMENT 97044478
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 330.

FEATURES
source

1. 411


```

1      FILING DATE: 7-JUNE-1995
2
3      ATTORNEY/AGENT INFORMATION:
4
5      NAME: Barry J. Swanson
6
7      REGISTRATION NUMBER: 33,215
8
9      REFERENCE/DOCKET NUMBER: NEX66
10
11     TELECOMMUNICATION INFORMATION:
12
13     TELEPHONE: (303) 268-0066
14
15     TELEFAX: (303) 268-0065
16
17     INFORMATION FOR SEQ ID NO: 14:
18
19     SEQUENCE CHARACTERISTICS:
20
21     LENGTH: 86 base pairs
22
23     TYPE: nucleic acid
24
25     STRANDEDNESS: single
26
27     TOPOLOGY: linear
28
29     MOLECULE TYPE: DNA
30
31     SEQUENCE DESCRIPTION: SEQ ID NO: 14:
32
33     US-08-991-743C-14

```

Query Match	47.9%	Score 18.2	DB 4	Length 86
Best Local Similarity	74.2%	Pred. No.26		
Matches 23; Conservative	0	Mismatches 8	Indels 0	Gaps 0

```
Qy 1 tgcctagaggaattcgaagtgcacatactt 31
    ||||| ||| | ||||| ||| |
Db 70 TGCTCAAGTGGGACTAGAGTGCAGGCAT 40
```

RESULT 15
 PCT-US96-08014-103/C
 ? Sequence 103, Application PC/TUS9608014
 ? GENERAL INFORMATION:
 ? APPLICANT: LARRY GOLD; NEBOJSA JANJIC; STEVEN RINGQUIST;
 ? APPLICANT: PAGATIS; PENELOPE J. TOOTHMAN
 ? TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
 ? TITLE OF INVENTION: LIGANDS TO TRANSFORMING GROWTH
 ? TITLE OF INVENTION: FACTOR (TGF) / PLATELET-DERIVED
 ? TITLE OF INVENTION: GROWTH FACTOR (PDGF) AND HUMAN
 ? TITLE OF INVENTION: KERATINOCYTE GROWTH FACTOR (hKGF)
 ? NUMBER OF SEQUENCES: 304
 ? CORRESPONDENCE ADDRESSES:
 ? ADDRESSEE: Swanson & Bratschun, L.L.C.
 ? STREET: 8400 E. Prentice Avenue, Suite 200
 ? City: Englewood
 ? STATE: Colorado
 ? COUNTRY: USA
 ? ZIP: 80111
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
 ? COMPUTER: IBM compatible
 ? OPERATING SYSTEM: MS-DOS
 ? SOFTWARE: Wordperfect 6.1
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: PCT/US96/08014
 ? FILING DATE:
 ? CLASSIFICATION:
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: 08/458,423
 ? FILING DATE: 02-JUNE-1995
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: 08/458,424
 ? FILING DATE: 02-JUNE-1995
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: 08/465,594
 ? FILING DATE: 05-JUNE-1995
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: 08/465,591
 ? FILING DATE: 05-JUNE-1995
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: 08/479,725
 ? FILING DATE: 07-JUNE-1995
 ? PRIOR APPLICATION DATA:
 ? APPLICATION NUMBER: 08/479,783

1 FILING DATE: 07-JUNE-1995
 2 PRIOR APPLICATION DATA:
 3 APPLICATION NUMBER: 08/618,693
 4 FILING DATE: 20-MARCH-1996
 5 ATTORNEY/AGENT INFORMATION:
 6 NAME: Barry J. Swanson
 7 REGISTRATION NUMBER: 33,215
 8 REFERENCE/DOCKET NUMBER:
 9 TELECOMMUNICATION INFORMATION:
 10 TELEPHONE: (303) 793-3333
 11 TELEFAX: (303) 793-3433
 12 INFORMATION FOR SEQ ID NO: 103:
 13 SEQUENCE CHARACTERISTICS:
 14 LENGTH: 86 base pairs
 15 TYPE: nucleic acid
 16 STRANDEDNESS: single
 17 TOPOLOGY: linear
 18 MOLECULE TYPE: DNA
 19 CBT-US96-08014-103

Query Match	47.98;	Score 18.2;	DB 5;	Length 86;
Best Local Similarity	74.28;	Pred. No. 26;		
Matches 23; Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

```

QY      1  tgcctagaggaattcgaagtacatact 31
          |||||  ||  ||  |||||  ||  ||
Db      70  TGCCTCACTGGGACTAGAACTGACGAGGCAT 40

```

Search completed: December 4, 2001, 18:58:38
Job time: 9264 sec

COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,693
FILING DATE: 20 MARCH 1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,783
FILING DATE: 7-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX42/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 86 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-618-693-14

Query Match 47.9%; Score 18.2; DB 1; Length 86;
Best Local Similarity 74.2%; Pred. No. 26;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1 tgcctagagaattcgagtgacgatactt 31
Db 70 TGCTCAAGTGGAGCTAGAGTGACGAGCAT 40

RESULT 13
US-08-973-124-103/C
Sequence 103, Application US/08973124
Patent No. 6207816
GENERAL INFORMATION:
APPLICANT: LARRY GOLD et al.
TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
TITLE OF INVENTION: LIGANDS TO GROWTH
NUMBER OF SEQUENCES: 304
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,124
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08014
FILING DATE: 30-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,423
FILING DATE: 02-JUNE-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/458,424
FILING DATE: 02-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,594
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,591
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,725
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,783
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,693
FILING DATE: 20-MARCH-1996
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 86 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-973-124-103

Query Match 47.9%; Score 18.2; DB 4; Length 86;
Best Local Similarity 74.2%; Pred. No. 26;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1 tgcctagagaattcgagtgacgatactt 31
Db 70 TGCTCAAGTGGAGCTAGAGTGACGAGCAT 40

RESULT 14
US-08-991-743C-14/C
Sequence 14, Application US/08991743C
Patent No. 6229002
GENERAL INFORMATION:
APPLICANT: NEBOJSA JANJIC, LARRY GOLD
TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR (PDGF) NUCLEIC
ACID LIGAND COMPLEXES
NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson and Bratschun, L.L.C.
STREET: 1745 Shea Center Drive, Suite 330
CITY: Highlands Ranch
STATE: Colorado
COUNTRY: USA
ZIP: 80129
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,743C
FILING DATE: 16-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,693
FILING DATE: 20-MARCH-1996
APPLICATION NUMBER: 08/479,783
FILING DATE: 7-JUNE-1995
APPLICATION NUMBER: 08/479,725


```

1  ATTORNEY/AGENT INFORMATION:
2  NAME: Tamarlin Ph.D., Lisa A.
3  REGISTRATION NUMBER: P-38,347
4  REFERENCE/DOCKET NUMBER: PD-3602
5  TELECOMMUNICATION INFORMATION:
6  TELEPHONE: (619) 455-5100
7  TELEFAX: (619) 455-5110
8  INFORMATION FOR SEQ ID NO: 1:
9  SEQUENCE CHARACTERISTICS:
10  LENGTH: 1991 base pairs
11  TYPE: nucleic acid
12  STRANDEDNESS: single
13  TOPOLOGY: linear
14  MOLECULE TYPE: DNA (genomic)
15  IMMEDIATE SOURCE:
16  CLONE: OmpL2
17  FEATURE:
18  NAME/KEY: CDS
19  LOCATION: 96..1715
20  OS-08-249-013-1

```

Query Match	48.4%	Score 18.4	DB 1	Length 1991
Best Local Similarity	69.4%	Pred. No. 41		
Matches 25; Conservative	0	Mismatches 11	Indels 0	Gaps 0

```

Qy      3  cctagaggaattcgaaagtcagatactttgycgc 38
          |||  |||  |||  |||  |||  |||  |||  |||
Db      1570 cTCAAGGGGcCTTAGAATTGGCGATCAATTGGGGC 1605

```

```

1      RESULT      8
2      US-08-886-863-1
3      : Sequence 1, Application US/08886863
4      : Patent No. 5824321
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: Haake, David A.
8      : TITLE OF INVENTION: CLONED leptospira OUTER MEMBRANE PROTEIN
9      : NUMBER OF SEQUENCES: 10
10     : CORRESPONDENCE ADDRESS:
11     : ADDRESSEE: Spensley Horn Jubas & Lubitz
12     : STREET: 1880 Century Park East, Suite 500
13     : CITY: Los Angeles
14     : STATE: California
15     : COUNTRY: USA
16     : ZIP: 90067
17     :
18     : COMPUTER READABLE FORM:
19     : MEDIUM TYPE: Floppy disk
20     : COMPUTER: IBM PC compatible
21     : OPERATING SYSTEM: PC-DOS/MS-DOS
22     : SOFTWARE: PatentIn Release #1.0, Version #1.25
23     : CURRENT APPLICATION DATA:
24     : APPLICATION NUMBER: US/08/886,863
25     : FILING DATE: 01-JUL-1997
26     : CLASSIFICATION: 435
27     :
28     : PRIOR APPLICATION DATA:
29     : APPLICATION NUMBER: US 08/249,013
30     : FILING DATE: 25-MAY-1994
31     : ATTORNEY/AGENT INFORMATION:
32     : NAME: Tumaarkin Ph.D., Lisa A.,
33     : REGISTRATION NUMBER: P-38,347
34     : REFERENCE/DOCKET NUMBER: PD-3602
35     : TELECOMMUNICATION INFORMATION:
36     : TELEPHONE: (619) 455-5100
37     : TELEFAX: (619) 455-5110
38     : INFORMATION FOR SEQ ID NO: 1:
39     : SEQUENCE CHARACTERISTICS:
40     : LENGTH: 1991 base pairs
41     : TYPE: nucleic acid
42     : STRANDEDNESS: single
43     : TOPOLOGY: linear
44     : MOLECULE TYPE: DNA (genomic)
45     : IMMEDIATE SOURCE:

```

```

; CLONE:  OmpL2
;
; FEATURE:
; NAME/KEY:  CDS
; LOCATION:  96..1715
;
US-08-886-863-1

```

Query Match	48.4%;	Score 18.4;	DB 1;	Length 1991;
Best Local Similarity	69.4%;	Pred. No. 41;		
Matches 25; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;

Qy	3	ctctagaaggaattcgaagtacgataactttggcgc	38
Db	1570	ctcAAGGGGccTTAGaATtGcGATCAATTTGGGGC	1605

```

RESULT 9
PCT-US95-06764-1
: Sequence 1, Application PC/TUS9506764
: GENERAL INFORMATION:
: APPLICANT: The Regents of the University of California
: TITLE OF INVENTION: CLONED leptospira OUTER MEMBRANE PROTEIN
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Fish & Richardson
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: California
: COUNTRY: USA
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/06764
: FILING DATE: 25-MAY-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Halle, Ph.D., Lisa A.,
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: FD3602
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 678-5070
: TELEFAX: (619) 678-5099
: INFORMATION FOR SEQ. ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1991 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: IMMEDIATE SOURCE:
: CLONE: Ompl2
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 96..1715
: PCT-US95-06764-1

Query Match 48.4%; Score 18.4; DB 5; Length 1991;
Best Local Similarity 69.4%; Pred. No. 41;
Matches 25; Conservative 0; Mismatches 11; Indels 0

DY 3 ctctagaagaattcgagaagtgaacatactttggcgc 38
||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1570 CTCAGGGGCGCTTAGAATTCGCATCAATTGGCGC 1605

RESULT 10
US-08-479-783A-14/C
Sequence 14, Application US/08479783A

```

OY	3	cctcagaggggaatttcgaagttagcaataacttttgcgagcc	38
Db	1570	CCTCAGGAGGGCCCTTTCGAATTTGCAGCATCAATTGGCGC	1605

RESULT 10
US-08-479-783A-14/c
; Sequence 14, Application US/08479783A

TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 11298 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: homo sapien
STRAIN: FcRI beta
US-07-869-933-31

Query Match 48.9%; Score 18.6; DB 1; Length 11298;
Best Local Similarity 72.7%; Pred. No. 49;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy 1 tgctctagagaattcgaatgcagatattt 33
Db 5630 TGTTCATAGCAATTGTAGTATGATGCTGTT 5662

RESULT 5
US-08-201-879A-2
Sequence 2, Application US/08201879A
Patent No. 5807988
GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
APPLICANT: JOUVIN, Marie-Helene
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/201,879A
FILING DATE: 24-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/869,933
FILING DATE: 16-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03419
FILING DATE: 16-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/234/NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 11298 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Homo sapiens

STRAIN: FcRI beta
FEATURE:
NAME/KEY: CDS
LOCATION: join(456..511, 1381..1510, 2026..2160, 4475..4531,
LOCATION: 5079..5237, 5640..5738, 7224..7319)
US-08-201-879A-2

Query Match 48.9%; Score 18.6; DB 1; Length 11298;
Best Local Similarity 72.7%; Pred. No. 49;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy 1 tgctctagagaattcgaatgcagatattt 33
Db 5630 TGTTCATAGCAATTGTAGTATGATGCTGTT 5662

RESULT 6
US-09-103-663-31
Sequence 31, Application US/09103663D
Patent No. 6171803
GENERAL INFORMATION:
APPLICANT: Kinet et al.
TITLE OF INVENTION: Isolation, characterization, and use of the human beta
TITLE OF INVENTION: subunit of the high affinity receptor for
TITLE OF INVENTION: immunoglobulin E.
FILE REFERENCE: 50490
CURRENT APPLICATION NUMBER: US/09/103,663D
CURRENT FILING DATE: 1998-06-23
EARLIER APPLICATION NUMBER: 07/869,933
EARLIER FILING DATE: 1992-04-16
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 31
LENGTH: 11298
TYPE: DNA
ORGANISM: Homo sapiens
US-09-103-663-31

Query Match 48.9%; Score 18.6; DB 4; Length 11298;
Best Local Similarity 72.7%; Pred. No. 49;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy 1 tgctctagagaattcgaatgcagatattt 33
Db 5630 tgctcaatagagaattcgaatgcagatattt 5662

RESULT 7
US-08-249-013-1
Sequence 1, Application US/08249013
Patent No. 5643754
GENERAL INFORMATION:
APPLICANT: Haake, David A.
TITLE OF INVENTION: CLONED leptospira OUTER MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Judas & Lubitz
STREET: 1880 Century Park East, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,013
FILING DATE: 25-MAY-1994
CLASSIFICATION: 435

Db 304 GCTCTGCAGATTTCATGTACCATCTTTTGCC 270

```
RESULT 2
US-08-458-356-163/C
; Sequence 163, Application US/08458356
; Patent No. 5942235
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,356
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,009
FILING DATE: 19-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066CHRTMS
INFORMATION FOR SEQ ID NO: 163:
SEQUENCE CHARACTERISTICS:
LENGTH: 598 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-458-356-163

Query Match 50.0%; Score 19; DB 2; Length 598;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 gctctagagaaatcgaatgcagatactttggc 36
||||| ||||| ||| ||||| ||||| ||||| |||||
Db 304 GCTCTGCAGATTTCATGTACCATCTTTTGCC 270

RESULT 3
US-08-460-736-163/C
; Sequence 163, Application US/08460736
; Patent No. 6265189
GENERAL INFORMATION:
APPLICANT: Paoletti, Enzo
APPLICANT: Tartaglia, James
APPLICANT: Cox, William I.
TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SEQUENCES: 217
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
```

```
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,736
FILING DATE: 02-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/184,009
FILING DATE: 19-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066CHRTMS
INFORMATION FOR SEQ ID NO: 163:
SEQUENCE CHARACTERISTICS:
LENGTH: 598 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-460-736-163

Query Match 50.0%; Score 19; DB 4; Length 598;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 gctctagagaaatcgaatgcagatactttggc 36
||||| ||||| ||| ||||| ||||| ||||| |||||
Db 304 GCTCTGCAGATTTCATGTACCATCTTTTGCC 270

RESULT 4
US-07-869-933-31
; Sequence 31, Application US/07869933
; Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: Kinet, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/154 NIH
TELECOMMUNICATION INFORMATION:
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:58:37 ; Search time 152.61 Seconds (with about 315 elements)

Title: US-09-915-060-7

Sequence: 1 **tgctctagaggaattcgaagtcacgatactt**tgcygc 38

Scoring table: IDENTITY_NUC

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
------------------	---------------	----

Listing first 45 summaries

```
Database :      Issued_Patents_NA: *
1  /acn3 6/atadata 7/4/na/5A COMP 000:*
```

```

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
C 1	19	50.0	598	2	US-08-184-009-163	Sequence 163, Appl	
C 2	19	50.0	598	2	US-08-438-356-163	Sequence 163, Appl	
C 3	19	50.0	598	4	US-08-460-736-163	Sequence 163, Appl	
C 4	18.6	48.9	11298	1	US-07-869-933-31	Sequence 31, Appl	
5	18.6	48.9	11298	1	US-08-201-879A-2	Sequence 2, Appl	
6	18.6	48.9	11298	4	US-09-103-663-31	Sequence 31, Appl	
7	18.4	48.4	11991	1	US-08-249-013-1	Sequence 1, Appl	
8	18.4	48.4	1991	1	US-08-886-863-1	Sequence 1, Appl	
9	18.4	48.4	1991	5	PCT-US95-06764-1	Sequence 1, Appl	
C 10	18.2	47.9	86	1	US-08-479-783A-14	Sequence 14, Appl	
C 11	18.2	47.9	86	1	US-08-479-725-14	Sequence 14, Appl	
C 12	18.2	47.9	86	1	US-08-618-693-14	Sequence 14, Appl	
C 13	18.2	47.9	86	4	US-08-973-124-103	Sequence 103, Appl	
C 14	18.2	47.9	86	4	US-08-991-743C-14	Sequence 14, Appl	
C 15	18.2	47.9	86	5	PCT-US96-08014-103	Sequence 103, Appl	
16	18	47.4	2236	2	US-08-484-993B-13	Sequence 13, Appl	
17	18	47.4	2236	2	US-08-484-158B-13	Sequence 13, Appl	
18	18	47.4	2236	2	US-08-484-596A-13	Sequence 13, Appl	
19	18	47.4	2236	2	US-08-480-150A-13	Sequence 13, Appl	
20	18	47.4	2236	3	US-08-458-731-13	Sequence 13, Appl	
21	18	47.4	2236	3	US-08-149-223A-13	Sequence 13, Appl	
22	17.8	46.8	3205	4	US-09-061-768A-3	Sequence 3, Appl	
23	17.8	46.8	4060	1	US-08-164-292B-1	Sequence 1, Appl	
24	17.8	46.8	4060	1	US-08-164-292B-3	Sequence 3, Appl	
25	17.8	46.8	4060	1	US-08-164-292B-5	Sequence 5, Appl	
26	17.8	46.8	4060	1	US-08-164-292B-7	Sequence 7, Appl	
27	17.8	46.8	4060	3	US-08-845-623-1	Sequence 1, Appl	

28	17.8	46.8	4060	3	US-08-845-622-3	Sequence 3, April 1
29	17.8	46.8	4060	3	US-08-845-623-5	Sequence 5, April 1
30	17.8	46.8	4060	3	US-08-845-623-7	Sequence 7, April 1
31	17.8	46.8	4060	3	US-08-815-927-1	Sequence 1, April 1
32	17.8	46.8	4060	3	US-08-815-927-3	Sequence 3, April 1
33	17.8	46.8	4060	3	US-08-815-927-5	Sequence 5, April 1
34	17.8	46.8	4060	3	US-08-815-927-7	Sequence 7, April 1
35	17.6	46.3	305	1	US-08-170-294-5	Sequence 5, April 1
36	17.6	46.3	305	1	US-08-664-854-5	Sequence 5, April 1
37	17.6	46.3	305	1	US-09-049-289-5	Sequence 5, April 1
38	17.6	46.3	2042	2	US-08-933-821-16	Sequence 16, April 1
39	17.6	46.3	2042	3	US-08-934-499-5	Sequence 5, April 1
40	17.6	46.3	2042	3	US-08-960-507-16	Sequence 16, April 1
41	17.6	46.3	2042	3	US-09-143-066-5	Sequence 5, April 1
42	17.6	46.3	2493	3	US-08-845-055-3	Sequence 3, April 1
43	17.6	46.3	3827	1	US-08-170-294-6	Sequence 6, April 1
44	17.6	46.3	3827	2	US-08-664-854-6	Sequence 6, April 1
45	17.6	46.3	3827	2	US-08-718-751-1	Sequence 1, April 1

ALIGNMENTS

RESULT 1
US-08-184-009-163/c
; Sequence 163, Application US/08184009

GENERAL INFORMATION:

APPLICANT: Tartagjia, James

TITLE OF INVENTION;	RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SENTENCES:	317

CORRESPONDENCE ADDRESS:

STREET: 530 Fifth Avenue
CITY: New York

STATE: NY
COUNTRY: USA

ZIP: 10036

MEDIUM TYPE: Floppy disk

```

;
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: DataTrn Polosco #1.0 Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: IIS/08/184 009

```

FILING DATE: 19-JAN-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Frommer William S

REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2530

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712
TELEEX: 425066CITBMS

SEQUENCE CHARACTERISTICS:

LENGTH: 598 base pairs
TYPE: nucleic acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear

```

MOLECULE TYPE: CDNA
 HIS-08-184-009-163

Query Match	50.0%;	Score 19;	DB 2;	Length 598;
-------------	--------	-----------	-------	-------------

Matches	25;	Conservative	0;	Mismatches	10;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	-----	--------	----	------	----

2 gctctagaggaattcgaagtgcacatacttctgc 36

XX	Borrelia burgdorferi; spirochete;bacterium; pathogen; Lyme disease;
KW	Epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW	Infection; diagnosis; characterisation; detection; ds.
XX	
OS	Borrelia burgdorferi.
PN	WO9858943-A1.
XX	
PD	30-DEC-1998.
XX	
PF	18-JUN-1998; 98WO-US12764.
XX	
PR	03-SEP-1997; 97US-0057483.
PR	20-JUN-1997; 97US-0050359.
PR	22-JUL-1997; 97US-0053344.
PR	22-JUL-1997; 97US-0053377.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PA	(MEDI-) MEDIMUNE INC.
PI	Clayton R, Dougherty BA, Fraser C, Iathigra R, Smith HO;
PI	White OR;
DR	WPI: 1999-081217/07.
XX	
PT	New isolated Borrelia burgdorferi nucleic acids - used to develop
PT	products for the detection, diagnosis, characterisation, prevention
PT	and therapy of infections, particularly Lyme disease
XX	
PS	Claim 1: Page 1023-1024; 1128pp; English.
XX	
CC	AAV20248 to AAV20402 represent polynucleotide sequences isolated from
CC	Borrelia burgdorferi (Bb). Products derived from Bb can be used for
CC	the detection, diagnosis, characterisation, prevention and therapy of
CC	Bb infections, e.g. Lyme disease. They can also be used for the
CC	production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC	to a family of motile, spiral-shaped bacteria called Spirochetes.
CC	Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC	endemic relapsing fever, and Lyme borreliosis, more commonly known as
CC	Lyme disease.
XX	
SQ	Sequence 2532 BP; 874 A; 451 C; 250 G; 955 T; 2 other;
XX	
Query Match:	50.5%; Score 19.2; DB 20; Length 2532;
Best Local Similarity	75.0%; Pred. No. 70;
Matches 24: Conservative	0; Mismatches 8; Indels 0; Gaps 0.
OY	3 ctctagagaattcgaagtacgaccttgg 34
Db	178 CTCTAGACGTACAATTATGTAAGAATCACTTTAG 147
RESULT 11	
ID	AAV20268/C
XX	AAV20268 standard; DNA; 5805 BP.
XX	
AC	AAV20268;
XX	
DT	04-MAY-1999 (first entry)
XX	
DE	Borrelia burgdorferi polynucleotide sequence #21.
XX	
KW	Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW	Epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW	Infection; diagnosis; characterisation; detection; ds.
OS	Borrelia burgdorferi.
XX	
PN	WO9858943-A1.
XX	
PD	30-DEC-1998.
XX	

XX	PR	18-JUN-1998;	98WO-US12764.
XX	PR	03-SEP-1997;	97US-0057483.
XX	PR	20-JUN-1997;	97US-0050359.
XX	PR	22-JUL-1997;	97US-0053344.
XX	PR	22-JUL-1997;	97US-0053377.
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	PA	(MEDI-) MEDIMUNE INC.	
XX	PI	Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO:	
XX	PI	White OR:	
XX	DR	WPI; 1999-081217/07.	
XX	PT	New isolated <i>Borrelia burgdorferi</i> nucleic acids - used to develop	
XX	PT	products for the detection, diagnosis, characterisation, prevention	
XX	PT	and therapy of infections, particularly Lyme disease	
XX	XX	Claim 1; Page 959-962; 1128pp; English.	
XX	XX	AA020248 to AA020402 represent polynucleotide sequences isolated from	
XX	CC	<i>Borrelia burgdorferi</i> (Bb). Products derived from Bb can be used for	
XX	CC	the detection, diagnosis, characterisation, prevention and therapy of	
XX	CC	Bb infections, e.g. Lyme disease. They can also be used for the	
XX	CC	production of biosynthetic products, e.g. enzymes. <i>Borrelia</i> belongs	
XX	CC	to a family of motile, spiral-shaped bacteria called Spirochetes.	
XX	CC	Spirochetes are pathogenic in humans and <i>Borrelia</i> causes epidemic and	
XX	CC	endemic relapsing fever, and Lyme borreliosis, more commonly known as	
XX	CC	Lyme disease.	
XX	SO	Sequence 5805 BP; 1878 A; 1000 C; 614 G; 2290 T; 23 other;	
XX	SO		
XX	XX	Query Match	50.5%; Score 19.2; DB 20; Length 5805;
XX	XX	Best Local Similarity	75.0%; Pred. No. 81;
XX	XX	Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;	
XX	Oy	3 ctctagaggaaatcgaagtgaacgactcttgg 34	
XX	Oy		
XX	Dd	1242 CTCCTAGACTACCAATTGAGTGAAGATCTTTTG 1211	
XX	RESULT 12		
XX	ID	AA020252/c	
XX	AC	AA020252 standard; DNA; 35515 BP.	
XX	AC	AA020252;	
XX	DT	04-MAY-1999 (first entry)	
XX	DE	<i>Borrelia burgdorferi</i> polynucleotide sequence #5.	
XX	KW	<i>Borrelia burgdorferi</i> ; spirochete; bacterium; pathogen; Lyme disease;	
XX	KW	epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;	
XX	KW	infection; diagnosis; characterisation; detection; ds.	
XX	OS	<i>Borrelia burgdorferi</i> .	
XX	PN	W09858943-A1.	
XX	PN		
XX	PD	30-DEC-1998.	
XX	PF	18-JUN-1998; 98WO-US12764.	
XX	PR	03-SEP-1997; 97US-0057483.	
XX	PR	20-JUN-1997; 97US-0050359.	
XX	PR	22-JUL-1997; 97US-0053344.	
XX	PR	22-JUL-1997; 97US-0053377.	
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	PA	(MEDI-) MEDIMUNE INC.	

AC AAC93671;
XX
DT 19-FEB-2001 (first entry)
XX
DE Cat flea HNC CLB2A cDNA complement, SEQ ID NO:164.
XX
KM Cat flea: head and nerve cord nucleic acid; HNC: flea infestation;
XX vaccine: antiparasitic; therapeutic target; diagnosis; detection; ss.
XX Ctenocephalides felis.
OS
PN W0200061621-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000MO-US09437.
XX
PR 09-APR-1999; 99US-0128704.
XX
PA (HESK-) HESKA CORP.
PI Brandt KS, Gaines PJ, Stinchcomb DR, Wisniewski N;
XX WPI: 2000-656323/63.
DR P-PSDB; AAB29618.
XX
PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic
PT acids useful for the prevention, diagnosis and treatment of flea
PT infestations -
XX
XX Claim 1; Page 298; 964pp; English.
XX
CC The invention relates to novel cat flea (Ctenocephalides felis) nucleic
CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
CC or head and nerve cord (HNC) tissue. The invention also relates to the
CC encoded proteins. The invention additionally encompasses expression
CC constructs, recombinant viruses and recombinant cells comprising the
CC nucleic acids of the invention, recombinant production of the proteins,
CC antibodies against the proteins, a method of identifying inhibitors of
CC the proteins, and compositions comprising the inhibitors for
CC administration to an animal. The nucleic acids, and the proteins they
CC encode may be used in the prevention, treatment and diagnosis of diseases
CC associated with flea infestations. For example, the nucleic acids may be
CC used to produce an HMT or HNC protein according to standard recombinant
CC DNA methodology by inserting the nucleic acids into a host cell and
CC culturing the cell to express the protein. The HMT and HNC nucleic acids
CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC and quantitate the presence of cat flea or other homologous nucleic acid
CC sequences in samples. They may also be used to study the expression and
CC function of the proteins and their role in metabolism. The HMT and HNC
CC proteins may be used as antigens in the production of specific
CC antibodies, and in assays to identify modulators (agonists and
CC antagonists) of HMT and/or HNC protein expression and activity. The
CC anti-HMT/HNC protein antibodies and antagonists may also be used to
CC downregulate protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of flea polypeptides
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC present sequence represents a cat flea HNC cDNA of the invention.
XX
SQ Sequence 631 BP; 150 A; 144 C; 82 G; 255 T; 0 other;

Query Match 50.5%; Score 19.2; DB 21; Length 631;
Best Local Similarity 75.0%; Pred. No. 55;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 6 tagaggaattcgaaatgacactcttgagc 37
||||| ||| ||| ||| ||| ||| ||| |||
Db 581 TAGAAGAATCAAAATGATGATATCTCTGTCG 550

RESULT 9
AAAX13016/c
1

ID AAX13016 standard; DNA: 2213 BP.
XX
AC AAX13016;
XX
DT 19-MAR-1999 (first entry)
XX
DE Enterococcus faecalis genome contig SEQ ID NO:79.
XX
KM Enterococcus faecalis; contig; detection; Enterococcal infection;
XX vaccine; attenuation; computer readable medium; ds.
XX
OS Enterococcus faecalis.
XX
PN W09850555-A2.
XX
PD 12-NOV-1998.
XX
PF 04-MAY-1998; 98MO-US08985.
XX
PR 14-NOV-1997; 97US-0066009.
PR 06-MAY-1997; 97US-0044031.
PR 16-MAY-1997; 97US-0046655.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Barash SC, Dillon PJ, Kunsch CA;
XX WPI: 1999-045171/04.
DR
XX
XX New Isolated Enterococcus faecalis polynucleotides and polypeptides
PT - used to develop products for the detection of Enterococcus and for
PT use in vaccines for prevention or attenuation of Enterococcus
PT infection.
XX
PS Claim 1; Page 558-559; 2084pp; English.
XX
CC A computer readable medium has been developed which has recorded on it
CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
CC AAX12398 to AAX13919 represent these nucleotide sequences which are
CC primary nucleotide sequences, also known as contigs. The computer-based
CC system can identify fragments of the Enterococcus faecalis genome with
CC commercial importance. The products can be used to detect the presence
CC of Enterococcus faecalis in samples. They can also be used for
CC diagnosing Enterococcal infection in an animal and monitoring
CC progression of disease, and for identifying agents which can be used to
CC modulate the growth or pathogenicity of Enterococcus faecalis, or
CC another related organism, in vivo or in vitro. In particular the
CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
CC can be used in vaccines to prevent or attenuate an Enterococcal
CC infection.
XX
SQ Sequence 2213 BP; 721 A; 342 C; 427 G; 720 T; 3 other;

Query Match 50.5%; Score 19.2; DB 20; Length 2213;
Best Local Similarity 75.0%; Pred. No. 68;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 4 tctaggaattcgaaatgacactcttgagc 35
||||| ||| ||| ||| ||| ||| ||| |||
Db 1008 TCTCGAGCAATTGGAAATGGCGATTCTTGTG 977

RESULT 10
AAX20295/c
ID AAX20295 standard; DNA: 2532 BP.
XX
AC AAX20295;
XX
DT 04-MAY-1999 (first entry)
XX
DE Borrelia burgdorferi polynucleotide sequence #48.
XX

OY 4 tctagagaattcgaatgcacatacttgc 34
||||| ||
DB 308 TCTAGAGTAACAAATTAGTGAAGACTTTTG 278

RESULT 6
AAA60982
ID AAA60982 standard; DNA: 1094 BP.

AC AAA60982;

DT 08-NOV-2000 (first entry)

DE Antibody nucleotide sequence.

KW Lida-protein; LDP, Lida-chromophore; LDC; antibody; lidamycin;
KM antibiotic; anticancer; tumour; cancer; ds.

OS Synthetic.

PN CNI251840-A.

PD 03-MAY-2000.

PF 13-OCT-1999; 99CN-0121668.

PR 13-OCT-1999; 99CN-0121668.

PA (MED1-) INST MEDICINAL BIOTECHNOLOGY CHINESE ACA.

PI Zhen Y, Li S, Jiang M;

DR WPI; 2000-432052/38.

PT Constitutive fusion protein of ridamycin as anticancer antibiotic and
single-chain antibody -

PS Disclosure; Page 2-3; 18pp; Chinese.

CC The present invention describes a fusion protein, designated lidamycin,
composed of lida-protein (LDP) and lida-chromophore (LDC), which acts as
an anticancer antibiotic and single-chain antibody. LDP and LDC are
joined via non-covalent bonds and can be splitted and recombined. A
single-chain antibody scfv can specifically join with IV-type collagenase
and can suppress its activity and in order to increase the permeability
of medicine to capillary tubes and the penetrability to real tumour, DNA
recombination and molecular recombination are used to prepare new-type
constitutive fusion protein LDM-FV of anticancer targeting medicine
lidamycin and single-chain antibody. Its molecular weight is about 37KDa.
CC It has the activity to inhibit IV-type collagenase and strong intrusion-
resistance kill action to cancer cells. It may be an ideal clinic
medicine. The present sequence represents an antibody nucleotide sequence
given in the exemplification of the present invention.

CC
XX Sequence 1094 BP; 230 A; 312 C; 325 G; 227 T; 0 other;

Query Match 51.1%; Score 19.4; DB 21; Length 1094;

Best Local Similarity 70.3%; Pred. No. 50;

Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 tgcctagaagaattcgaatgcacatacttgcg 37
||| ||| | ||||| ||||| |||||
DB 460 tctcgcagagccagtgaaagtgttcatactatgcgc 496

RESULT 7

ID AAC93670 standard; cDNA: 631 BP.

AC AAC93670;

XX

DT 19-FEB-2001 (first entry)

DE Cat flea HNC CLBP2A cDNA, SEQ ID NO:162.

KW Cat flea; head and nerve cord nucleic acid; HNC; flea infestation;
KM vaccine; antiparasitic; therapeutic target; diagnosis; detection; ss.

XX Ctenocephalides felis.

PN WO200061621-A2.

PD 19-OCT-2000.

PF 07-APR-2000; 2000WO-US09437.

PR 09-APR-1999; 99US-0128704.

PA (HESK-) HESKA CORP.

PI Brandt KS, Gaines PJ, Stinchcomb DF, Wisniewski N;

DR WPI; 2000-656323/63.

DR P-PSTDB; AAB29618.

PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic
acids useful for the prevention, diagnosis and treatment of flea
infestations -

PS Claim 1; Page 296-297; 964pp; English.

CC The invention relates to novel cat flea (Ctenocephalides felis) nucleic
acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
or head and nerve cord (HNC) tissue. The invention also relates to the
CC encoded proteins. The invention additionally encompasses expression
CC constructs, recombinant viruses and recombinant cells comprising the
CC nucleic acids of the invention, recombinant production of the proteins,
CC antibodies against the proteins, a method of identifying inhibitors of
CC the proteins, and compositions comprising the inhibitors for
CC administration to an animal. The nucleic acids, and the proteins they
CC encode may be used in the prevention, treatment and diagnosis of diseases
CC associated with flea infestations. For example, the nucleic acids may be
CC used to produce an HMT or HNC protein according to standard recombinant
CC DNA methodology by inserting the nucleic acids into a host cell and
CC culturing the cell to express the protein. The HMT and HNC nucleic acids
CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC and quantitate the presence of cat flea or other homologous nucleic acid
CC sequences in samples. They may also be used to study the expression and
CC function of the proteins and their role in metabolism. The HMT and HNC
CC proteins may be used as antigens in the production of specific
CC antibodies, and in assays to identify modulators (agonists and
CC antagonists) of HMT and/or HNC protein expression and activity. The
CC anti-HMT/HNC protein antibodies and antagonists may also be used to
CC downregulate protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of flea polypeptides
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC present sequence represents a cat flea HNC cDNA of the invention.

CC
XX Sequence 631 BP; 255 A; 82 C; 144 G; 150 T; 0 other;

Query Match 50.5%; Score 19.2; DB 21; Length 631;

Best Local Similarity 75.0%; Pred. No. 55;

Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 6 tagagaattcgaatgcacatacttgcg 37
||| ||| || || ||| ||||| |||||
DB 51 tagaagaatacaaatgaagatacttcttgcg 82

RESULT 8

ID AAC93671/c standard; cDNA: 631 BP.

XX

```
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150684.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
```

```
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 52.6%; Score 20; DB 21; Length 1611;
Best Local Similarity 72.2%; Pred. No. 29;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Oy 2 gctcagsgaatcgaagtcagatcttggcg 37
Db 998 GCACTAATATATAPCCGATGGAAGATGTTTGGCG 963
```

RESULT 5
ID AAX20344/c
AAX20344; standard: DNA; 976 BP.
AAX20344;
04-MAY-1999 (first entry)

Borrelia burgdorferi polynucleotide sequence #97.
Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
infection; diagnosis; characterisation; detection; ds.

Borrelia burgdorferi.

W09858943-A1.

30-DEC-1998.

18-JUN-1998; 98MO-US12764.

03-SEP-1997; 97US-0057483.

20-JUN-1997; 97US-0050359.

22-JUL-1997; 97US-0053344.

22-JUL-1997; 97US-0053377.

(HUMA-) HUMAN GENOME SCI INC.
(MEDI-) MEDIMUNE INC.

Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
White OR;
WPI. 1999-081217/07.

New isolated Borrelia burgdorferi nucleic acids - used to develop
products for the detection, diagnosis, characterisation, prevention
and therapy of infections, particularly Lyme disease

Claim 1: Page 1077-1078; 1128pp; English.

AAX20248 to AAX20402 represent polynucleotide sequences isolated from
Borrelia burgdorferi (Bb). Products derived from Bb can be used for
the detection, diagnosis, characterisation, prevention and therapy of
Bb infections, e.g. Lyme disease. They can also be used for the
production of biosynthetic products, e.g. enzymes. Borrelia belongs
to a family of motile, spiral-shaped bacteria called Spirochetes.
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and
endemic relapsing fever, and Lyme borreliosis, more commonly known as
Lyme disease.

Sequence 976 BP; 308 A; 171 C; 94 G; 400 T; 3 other;

Query Match 52.1%; Score 19.8; DB 20; Length 976;
Best Local Similarity 77.4%; Pred. No. 33;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Caps 0;

CC morphogenesis, recombination, metabolic or catabolic pathway
CC engineering. Using ESTs provides several advantages over genomic or
CC random cDNA clones including elimination of redundancy as one spot on an
CC array equals one gene or open reading frame, and organisation of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC *Fusarium venenatum*; AAF11248 to AAF11853 represents ESTs from
CC *Aspergillus niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and
CC AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are
CC all specifically claimed in the present invention.
XX
SQ Sequence 729 BP; 173 A; 194 C; 187 G; 175 T; 0 other;

Query Match 53.7%; Score 20.4; DB 21; Length 729;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 6 tagaggaatcgaagtacgatactttg 35
||||||| | ||||| ||| ||
Db 653 TAGAGGAATTGAGCTGACGAACTCCGCG 624

RESULT 4
AAC42471/c
ID AAC42471 standard; DNA; 1611 BP.
XX
AC AAC42471;
XX
DT 17-OCN-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35686.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145109.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145917.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

XX Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
XX Disclosure; Page 32-33; 57pp; English.
XX
CC The PITSIRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the human PITSIRE protein kinase isoform
CC alpha2-2. This sequence contains the IRES. The IRES sequence and
CC fragments of it may be used to induce cell cycle dependent initiation
CC of translation in eukaryotic cells. Vectors containing the IRES may be
CC used for the preparation of compositions for the treatment of and/or
CC prevention of a disease by gene therapy. Such diseases may be cancer
CC and restenosis.
CC
XX
SQ Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other;

Query Match 57.9%; Score 22; DB 21; Length 2471;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 17 gaagtcacatactttgagcgc 38
Db 7 gaagtcacatactttgagcgc 28

RESULT 2
AAAX20280
ID AAAX20280 standard; DNA; 3653 BP.
XX
AC AAAX20280;
XX
DT 04-MAY-1999 (first entry)
XX
DE Borrelia burgdorferi polynucleotide sequence #33.
XX
DE
XX
KW Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW infection; diagnosis; characterisation; detection; ds.
XX
OS Borrelia burgdorferi.
XX
PN MO9658943-A1.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 96MO-US12764.
XX
PR 03-SEP-1997; 97US-0057483.
PR 20-JUN-1997; 97US-0050359.
PR 22-JUL-1997; 97US-0053344.
PR 22-JUL-1997; 97US-0053377.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMUNE INC.
XX
XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
PI White OR;
XX
XX WPI; 1999-081217/07.
XX
PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of infections, particularly Lyme disease
XX
XX Claim 1: Page 994-996; 1128pp; English.
XX
XX AAAX20280 to AAAX20402 represent polynucleotide sequences isolated from
XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for
CC

CC the detection, diagnosis, characterisation, prevention and therapy of
CC Bb infections, e.g. Lyme disease. They can also be used for the
CC production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC to a family of motile, spiral-shaped bacteria called Spirochetes.
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as
CC Lyme disease.
XX
SQ Sequence 3653 BP; 1587 A; 348 C; 690 G; 1027 T; 1 other;

Query Match 54.7%; Score 20.8; DB 20; Length 3653;
Best Local Similarity 78.1%; Pred. No. 15;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 3 cctagaggaattcgaagtcacatactttg 34
Db 3220 ccttagagtaacaattagtcgaatactttg 3251

RESULT 3
AAFI1951/C
ID AAFI1951 standard; CDNA; 729 BP.
XX
AC AAFI1951;
XX
DT 13-MAR-2001 (first entry)
XX
DE Aspergillus oryzae EST SEQ ID NO:4474.
XX
DE
XX Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus oryzae.
XX
PN WO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000MO-US07781.
XX
PR 22-MAR-1999; 99US-0273623.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MW, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;
XX
XX WPI; 2000-594572/56.
XX
PT Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XX

Claim 88; Page 1910; 3161pp; English.
XX
XX The present invention describes a method for monitoring differential
XX expression of genes in a first filamentous fungal (FF) cell relative to
XX expression of the same genes in one or more second filamentous fungal
XX cells. The method uses fluorescence-labeled nucleic acids isolated from
XX the FF cells and a substrate of expressed sequence tags (EST). The ESTs
XX are used in the methods for monitoring differential expression of genes
XX in a first filamentous fungal (FF) cell relative to expression of the
XX same genes in one or more second filamentous fungal cells. Monitoring
XX the global expression of genes from FF cells allows the production
XX potential of the microorganisms to be improved. New genes may be
XX discovered, possible functions of unknown open reading frames can be
XX identified and gene copy number variation and stability can be
XX monitored. The expression of genes can be used to study how FF cells
XX adapt to changes in culture conditions, environmental stress, spore
CC

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 19:03:47 ; Search time 355.05 Seconds
(without alignments)
91.757 Million cell updates/sec

Title: US-09-915-060-7

Perfect score: 38
Sequence: 1 tgcctagaggaaatcgaagtgcacatacttggcgc 38

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_1101:*

1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	57.9	2471	21	AAA73711
2	20.8	54.7	3653	20	AAK20280
3	20.4	53.7	729	21	AAE11951
4	20	52.6	1611	21	AAK42471
5	19.8	52.1	976	20	AAK20344
6	19.4	51.1	1094	21	AAA60982
7	19.2	50.5	631	21	AAK93670
8	19.2	50.5	631	21	AAK93671
9	19.2	50.5	2213	20	AAK13016
10	19.2	50.5	2532	20	AAK20295
11	19.2	50.5	5805	20	AAK20268

C	12	19.2	50.5	35515	20	AAK20252
C	13	19	50.0	151	16	AAK22774
C	14	19	50.0	598	15	AAK67872
C	15	19	50.0	598	20	AAK08484
C	16	19	50.0	738	21	AAK14903
C	17	19	50.0	1046	21	AAK36690
C	18	19	50.0	2995	22	AAK58266
C	19	19	50.0	3419	21	AAK14902
C	20	18.8	49.5	40	21	AAK59884
C	21	18.8	49.5	570	22	AAK28978
C	22	18.8	49.5	858	20	AAK34757
C	23	18.8	49.5	1983	22	AAK25583
C	24	18.8	49.5	2097	19	AAK52027
C	25	18.8	49.5	2169	19	AAK13930
C	26	18.6	48.9	445	20	AAK1393
C	27	18.6	48.9	495	22	AAK12224
C	28	18.6	48.9	533	22	AAK68252
C	29	18.6	48.9	731	22	AAK03845
C	30	18.6	48.9	1090	21	AAK5565
C	31	18.6	48.9	1190	21	AAK54568
C	32	18.6	48.9	1191	22	AAK45561
C	33	18.6	48.9	1229	22	AAK45568
C	34	18.6	48.9	1248	21	AAK54275
C	35	18.6	48.9	3137	22	AAK17117
C	36	18.6	48.9	3455	20	AAK1385
C	37	18.6	48.9	4131	22	AAK77689
C	38	18.6	48.9	11298	18	AAK6756
C	39	18.6	48.9	11298	19	AAK5661
C	40	18.6	48.9	11298	21	AAK20937
C	41	18.6	48.9	11298	21	AAK34815
C	42	18.6	48.9	11298	22	AAK92144
C	43	18.6	48.9	11357	14	AAK51024
C	44	18.6	48.9	21742	21	AAK20938
C	45	18.6	48.9	21742	21	AAK34816

ALIGNMENTS

RESULT	1	
AAA73711		
ID	AAA73711	standard; DNA; 2471 BP.
XX		
AC	AAA73711;	
XX		
DT	07-DEC-2000	(first entry)
XX		
DE	Human PITSURE protein kinase isoform alpha2-2.	
XX		
KW	IRRS; internal ribosome entry site; PITSURE; human; gene therapy;	
XX	cancer; restenosis; p58; p110; protein kinase; ds.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	misc_signal	907..1128
FT		/*tag= a
FT		/label= IRRS
FT		/note= "internal ribosome entry site"
XX		
PN	WO200044896-A1.	
XX		
PD	03-AUG-2000.	
XX		
XX	26-JAN-2000; 2000WO-EP00643.	
PF		
XX	26-JAN-1999; 99EP-0200216.	
PR		
XX	(VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.	
PA		
XX	Cornelis S, Beyaert R;	
XX		
PI		
XX		
DR	WPI; 2000-499331/44.	

Borrelia burgdorfe
Human gene signatu
I3L/Mutine IFN-gam
Vaccinia I3L promo
Nucleotide sequenc
Arabidopsis thalia
Human polynucleoti
Nucleotide sequenc
Plasmodium yoelli
Drosophila melanog
H. pylori HP5084 e
Helicobacter polyP
H. pylori GHP 761
445 bp fragment of
Human cDNA clone (
Human lung tumour
Human cDNA clone (
Peroxidase related
Arabidopsis thalia
peroxidase gene.
PER4 peroxidase re
Arabidopsis thalia
Human cDNA sequenc
Nucleic acid encod
Human variant fcep
Human high affinit
Human beta subunit
Human high affinit
Human adenosine re
Human ICBR beta c
Human FCER1 beta c
Human high affinit
Human adenosine re


```
/product="PITSLRE isoform PBETA22"
/protein_id="AA19595.1"
/db_xref="GI:507429"
/translation="MGDEKDSKVKYTLIDELQEKRRKKEQEKAEIKRLKNSDRSK
RDLSEGLRDHCEMTITNSPYRRSDSMEDGDEDSLAIKPPQOMSRKEVHHKRD
EKREKKHARVKKRERERKRRHEEDOKARERERKRRMARHSHRRRGNGVCLF
RDRLQELERKRRERKRRKREKQEKREQEKREKREKREKREKREKREKREKREK
SDYKASHMRSRPPRRPRELGDGRKPYKEKMERDLSDLODISDEKRTSSAE
SSASGSGSGSEEEEEESEESSEESSEESSEESSEESSEESSEESSEESSEESSE
EBSSEDEERENENHLLVPRSRDRDROSGESEEAEVEEGESTPOSSALTBEDYPPDPA
LSPIELKQELPKYLPALQGRSVDERQCLNRIEEGYGVYRAKDKTDEIVALKRLK
MEKNEGEPTITSLREINTILKAQHPNIVYREIVVGSNMOKIYIVNMYEHDLSME
TMKOPFLPGEVKITLMQILRGVKKHLDNWLHRLDLKTNLLSHAGILKYGDFGLA
YGSPLKAVTPVYVYTWYRAPELLGAKEYSTAYDMMVSGCIEGELLORLPEFNGSEI
DQINKYKELGTSEKTPXSELPYKSKMTPSRHPYNNIRKRGALLISGCGDLMK
PLTFEGRRTISAEDGLKHEYFRETPIDPSMFPYMPAKSEQOVRKGTSPRPPEG
GYSOLGDDDLKETGEFHLITTTNOGASAGPFSLKF"

BASE COUNT      693 a      599 c      808 g      365 t
ORIGIN

Query Match      57.9%; Score 22; DB 9; Length 2465;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 gaagtgcacatactttggcgc 38
|||||
Db 7 GAAGTGACGATCTTTGGCGC 28

RESULT 12
AX03422 2471 bp DNA PAT 21-SEP-2000
LOCUS AX03422
DEFINITION Sequence 3 from Patent WO0044896.
ACCESSION AX03422
VERSION AX03422.1 GI:10280183
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2471)
AUTHORS Beyaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
JOURNAL Patent: WO 0044896-A 3 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE); BEYAERT RUDI (BE); CORNELIS
SIGRID (BE)
FEATURES
source location/Qualifiers
1..2471
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1..2471
/note="the IRES-activity containing sequence
PITSLRE protein kinase (p110pitslre) (isoform alfa 2-2)"
BASE COUNT      697 a      601 c      809 g      364 t
ORIGIN

Query Match      57.9%; Score 22; DB 6; Length 2471;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 gaagtgcacatactttggcgc 38
|||||
Db 7 GAAGTGACGATCTTTGGCGC 28

RESULT 13
HSU04816 2471 bp mRNA PRI 08-JUL-1994
LOCUS HSU04816
DEFINITION Human protein kinase PITSLRE alpha 2-2 mRNA, complete cds.
ACCESSION U04816
```

```
VERSION U04816.1 GI:507159
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2471)
AUTHORS Xiang,J., Lahfi,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE Molecular cloning and expression of alternatively spliced PITSLRE
protein kinase isoforms
JOURNAL J. Biol. Chem. 269, 15786-15794 (1994)
MEDLINE 94253170
REFERENCE 2 (bases 1 to 2471)
AUTHORS Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's
Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
Memphis, TN 38101, USA
FEATURES
source location/Qualifiers
1..2471
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1p36-2"
/cell_line="Hela S3"
/cell_type="epitheloid"
/tissue_type="cervix"
/clone_lib="Hela cDNA library"
/dev_stage="adult"
1..2471
112..2445
/codon_start=1
/product="PITSLRE alpha 2-2"
/protein_id="AA19582.1"
/db_xref="GI:507160"
/translation="MGDEKDSKVKYTLIDELQEKRRKKEQEKAEIKRLKNSDRSK
RDLSEGLRDHCEMTITNSPYRRSDSMEDGDEDSLAIKPPQOMSRKEVHHKRD
EKREKKHARVKKRERERKRRHEEDOKARERERKRRMARHSHRRRGNGVCLF
RDRLQELERKRRERKRRKREKQEKREQEKREKREKREKREKREKREKREKREKRE
SDYKASHMRSRPPRRPRELGDGRKPYKEKMERDLSDLODISDEKRTSSAE
SSASGSGSGSEEEEEESEESSEESSEESSEESSEESSEESSEESSEESSEESSE
EBSSEDEERENENHLLVPRSRDRDROSGESEEAEVEEGESTPOSSALTBEDYPPD
PALSPIELKQELPKYLPALQGRSVDERQCLNRIEEGYGVYRAKDKTDEIVALKR
LMKEKEGEPTITSLREINTILKAQHPNIVYREIVVGSNMOKIYIVNMYEHDLSME
TMKOPFLPGEVKITLMQILRGVKKHLDNWLHRLDLKTNLLSHAGILKYGDFGLA
REYGSPLKAVTPVYVYTWYRAPELLGAKEYSTAYDMMVSGCIEGELLORLPEFNGS
EIDQINKYKELGTSEKTPXSELPYKSKMTPSRHPYNNIRKRGALLISGCGDLMK
NKFITFGRRTISAEDGLKHEYFRETPIDPSMFPYMPAKSEQOVRKGTSPRPPEG
GLGYSOLGDDDLKETGEFHLITTTNOGASAGPFSLKF"

BASE COUNT      697 a      601 c      809 g      364 t
ORIGIN

Query Match      57.9%; Score 22; DB 9; Length 2471;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 gaagtgcacatactttggcgc 38
|||||
Db 7 GAAGTGACGATCTTTGGCGC 28

RESULT 14
HSU07704 2471 bp mRNA PRI 08-JUL-1994
LOCUS HSU07704
DEFINITION Human protein kinase PITSLRE isoform PBETA21 (p58CDC2L1) mRNA,
complete cds.
ACCESSION U07704
VERSION U07704.1 GI:507426
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```


gene
CDS
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2349
/gene="CDC21.1"
626..2323
/gene="CDC21.1"
/note="corresponding genomic sequence deposited as GenBank
Accession Numbers AF080678-Af080688, AF092429, and
AF092430"
/codon_start=1
/product="PITSLRE protein kinase alpha SV11 isoform"
/protein_id="AAC72082.1"
/db_xref="GI:3850314"
/translation="MREYSDKVKASHRSRPPRRREFLGDGRKPKREKMEERD
LISDIDSDSRKTSNAESSASGSGSEEEENEHLVPESTRDRSGSEAEED
BEETGNSSEASDAEVSSEEMSEDEEENEHLVPESTRDRSGSEAEED
VGGTPQSSALTEGDYVDPSPALSPIELKQPLPKPLQGRSVSEFQCLNRIEETG
GCVVYRAKDKTDEIVALKRLMEKEKEGPTITSLREINTILKAQHPNIVYREIVG
SNMDKIVYVMNVVEHDLKSLMETMKQPLPQEVKTIQMLRGVHLHDNMLHRDLK
TSNLLSHAGILKVGDFGLAREYGPLKATPYVYVTLMYRAPELLGAKKESTAVDMK
SYGCIIEGFLTKQPLPQKSEIDQNKYFKDLPSEKIMWGYSTLPAVKKMTSEHP
YNNLKRKRGALISDQEPDLANKFTYTFPGRIASADGLKHEFTREPLIDPSMPTW
PAKSEQKVKRSTSPREPGEGLGYSQLDDDLKETFHLTTTNOGASAGPGSLKF"

Query Match
Best Local Similarity 57.9%; Score 22; DB 9; Length 2349;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 17 gaagtgaacatactttggcgc 38
7 GAAGTGACGATCTTTGGCGC 28

RESULT 10
LOCUS HSU04817 2444 bp mRNA PRI 08-JUL-1994
DEFINITION Human protein kinase PITSLRE alpha 2-3 mRNA, complete cds.
ACCESSION U04817
VERSION U04817.1 GI:507161
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2444)
AUTHORS Xiang,J., Lahti,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE Molecular cloning and expression of alternatively spliced PITSLRE
protein kinase isoforms
JOURNAL J. Biol. Chem. 269, 15786-15794 (1994)
MEDLINE 94253170
REFERENCE
2 (bases 1 to 2444)
AUTHORS Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's
Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
Memphis, TN 38101, USA
FEATURES
source
1..2444
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1p36-2"
/cell_line="Hela S3"
/cell_type="epitheloid"
/tissue_type="cervix"
/clone_lib="Hela cDNA library"
/dev_stage="adult"
1..2444
mrna
CDS
112..2418

gene
CDS
/codon_start=1
/product="PITSLRE alpha 2-3"
/protein_id="AAA19583.1"
/db_xref="GI:507162"
/translation="MDEKDSWKVKTLDELTLQKRRKROEKEAETIKRLKNSDDRSDK
RDLSEGLDHCMEITIRNSPYRREDSDMDREEDDSIAIKPPQMSRREYVHNRKD
EKREKKHARVKKRREHRRKRREODKARREWRORREARREYSAHRTMREYDQKVAS
EKREKRRKREODKREODKREKREKREKREKREKREKREKREKREKREKREKREK
HMSRSPRRPPELGDGRKPKYKEMERDLISDIDSDSRKTSNAESSASG
SGSEEEEEEDEEBCSTSESEEEEEEDEEBCSTSESEEEEEEDEEBCSTSESEEE
ERENENHLVVPESRFRDROSGSESEAEVEEGTPOSALTEGDYVDPSPALSPIELK
QELPKYLPALQGRSVSEFQCLNRIEETGCVVYRAKDKTDEIVALKRLMEKEKEG
EPTSLREINTILKAQHPNIVYREIVGSMOKIYVMNVVEHDLKSLMETMKQPL
PQEVKTIQMLRGVHLHDNMLHRDLKTSNLLSHAGILKVGDFGLAREYGPLKA
YTPVYVTLMYRAPELLGAKKESTAVDMKSYGCIIEGFLTKQPLPQKSEIDQNKY
KDIQTPSEKIMWGYSTLPAVKKMTSEHPYNNLKRKRGALISDQEPDLANKFTYTFPG
RISAEGLKHEFTREPLIDPSMPTWPAKSEQKVKRSTSPREPGEGLGYSQLDDDLK
ETFLTTTNOGASAGPGSLKF"

Query Match
Best Local Similarity 57.9%; Score 22; DB 9; Length 2444;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 17 gaagtgaacatactttggcgc 38
7 GAAGTGACGATCTTTGGCGC 28

RESULT 11
LOCUS HSU07705 2465 bp mRNA PRI 08-JUL-1994
DEFINITION Human protein kinase PITSLRE isoform PBETA22 (p58CDC21.1) mRNA,
complete cds.
ACCESSION U07705
VERSION U07705.1 GI:507428
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2465)
AUTHORS Xiang,J., Lahti,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE Molecular cloning and expression of alternatively spliced PITSLRE
protein kinase isoforms
JOURNAL J. Biol. Chem. 269, 15786-15794 (1994)
MEDLINE 94253170
REFERENCE
2 (bases 1 to 2465)
AUTHORS Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-1994) Vincent J. Kidd, St. Jude Children's
Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
Memphis, TN 38101, USA
FEATURES
source
1..2465
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36-2"
/cell_line="Hela S3"
/cell_type="epitheloid"
/tissue_type="cervix"
/clone_lib="Hela cDNA library"
/dev_stage="adult"
1..2465
112..2439
/gene="p58CDC21.1"
112..2439
/gene="p58CDC21.1"
/function="protein kinase"
/codon_start=1

SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2226)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and MPM21-22 genes on human chromosome 1p36.3 and their linkage to D122
JOURNAL Genome Res. 8 (9), 929-939 (1998)
MEDLINE 98424414
REFERENCE 2 (bases 1 to 2226)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES
source
1. .2226
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1. .2226
/gene="CDC2L1"
620. .2200
/note="corresponding genomic sequence deposited as GenBank Accession Numbers AF080678-AF080688, AF092429, and AF092430"
/product="PITSLRE protein kinase alpha SV4 isoform"
/protein_id="AAC72078.1"
/db_xref="GI:3850306"
/translation="MEERDLISLDQDISSEKRTSSAESGSAAGSGSEEEEREE
EEEGSTSESEEESEEEETGSGSEASQSAEEVSEEMSEDEERENHLLV
VPESRFDSDGSESEEAEEVEEGTPOSSALTEGDPSPALSPIELKQELPYLAL
QGRSVPEFCNRIEETGYVYRAKDKTDEIVALKRLMEKEKEGFPITSLREIN
TILKQHPNIVTVREIYVSGNMKIYVMVYVHDLKSLMETKQEPLEGEVTLTIO
LLRGVRLHNDNMLHDLKTSNLLISAGILKYGDCIAGEGSPKATPPYVVTWY
RABPLLGAKEYSTAVDMMSVGCITGELTOKPLPEKSTIDINKYFKDLGTSKRI
WPGSELPAVKMTFSEHPYNNLKRREGALLSDQGDLMKFLTYFPGRRTSADGLK
HEYFRETPLIDPSMPTWPAKSEQOVRKRGTSRPEEGIGYSQLDDDLKETGAGHL
TTTNOGASAGAPFSLKE"

BASE COUNT 638 a 541 c 705 g 342 t
ORIGIN

Query Match 57.9%; Score 22; DB 9; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 gaagtacgatactttggcgc 38
|||||
Db 7 GAAGTGACGATACCTTTGGCGC 28

RESULT 8
LOCUS HS004818 2340 bp mRNA PRI 08-JUL-1994
DEFINITION Human protein kinase PITSLRE alpha 2-4 mRNA, complete cds.
ACCESSION U04818
VERSION U04818.1 GI:507163
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2340)
AUTHORS Xiang,J., Lahli,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE Molecular cloning and expression of alternatively spliced PITSLRE protein kinase isoforms
JOURNAL J. Biol. Chem. 269, 15786-15794 (1994)

MEDLINE 94253170
REFERENCE 2 (bases 1 to 2340)
AUTHORS Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St., Memphis, TN 38101, USA
FEATURES
source
1. .2340
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1p36-2"
/cell_line="Hela S3"
/cell_type="epitheloid"
/tissue_type="cervix"
/clone_lib="Hela CDNA library"
/dev_stage="adult"
1. .2340
/codon_start=1
626. .2314
/product="PITSLRE alpha 2-4"
/protein_id="AAA19584.1"
/db_xref="GI:507164"
/translation="MREDYSDKVKASHMSRPPRRREFELGGRKPKVEKMEERD
LLSDLDQDISSEKRTSSAESGSGSEEEERENHLLVPESRFDSDGSESEAEVEEG
ETGNSSEASQSAEEVSEEMSEDEERENHLLVPESRFDSDGSESEAEVEEG
GPSSALTEGDPVDPSPALSPIELKQELPYLALQGRSVPEECINRIEETGYV
YVRADKKTDEIVALKRLMEKEKEGFPITSLREINTILKQHPNIVTVREIYVSGNM
DITVVMNVVHDLKSLMETMQLPEEVVTLMIQLRGVHLHNDNMLHRDLKTSN
LLLSHAGILKVGDFGLAREYGSPLKAYTPVVTWYVRAPELLGAKEYSTAVDMMSV
CIFGELTOKPLPEPKSEIDQINKYFKDLGTSKIMPGSELPAVKMTFSRHPYNN
LKRREGALLSDQGDLMKFLTYFPGRRTSADGLKHEYFRETPLIDPSMPTWPAK
SEQOVRKRGTSRPEEGIGYSQLDDDLKETGAGHLTTTNOGASAGCGFSLKE"

BASE COUNT 666 a 576 c 741 g 357 t
ORIGIN

Query Match 57.9%; Score 22; DB 9; Length 2340;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 gaagtacgatactttggcgc 38
|||||
Db 7 GAAGTGACGATACCTTTGGCGC 28

RESULT 9
LOCUS AF067517 2349 bp mRNA PRI 07-NOV-1998
DEFINITION Homo sapiens PITSLRE protein kinase alpha SV11 isoform (CDC2L1)
ACCESSION AF067517
VERSION AF067517.1 GI:3850313
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2349)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and MPM21-22 genes on human chromosome 1p36.3 and their linkage to D122
JOURNAL Genome Res. 8 (9), 929-939 (1998)
MEDLINE 98424414
REFERENCE 2 (bases 1 to 2349)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES
source
1. .2349
Location/Qualifiers

```
repeat_region      /rpt_family="Alu"
                    6900..6913
repeat_region      /rpt_family="T-rich"
                    6914..7198
misc_feature        /rpt_family="Alu"
                    6928..6943
                    /note="similar to EST AA760985 (NID:92809915) nx32407.s1"
repeat_region      9140..9425
                    /rpt_family="Alu"
repeat_region      9426..9466
                    /rpt_family="(TAA)n"
repeat_region      9859..9991
                    /rpt_family="L1"
repeat_region      9992..10519
                    /rpt_family="MER1-type"
repeat_region      10520..11545
                    /rpt_family="L1"
                    complement(12187..12419)
STS                /db_xref="GI:1222793"
                    13007..13164
repeat_region      /rpt_family="MTR"
                    13312..13338
repeat_region      /rpt_family="AT-rich"
                    13905..13937
repeat_region      /rpt_family="(TTA)n"
                    13938..14185
repeat_region      /rpt_family="Alu"
                    14776..14800
repeat_region      /rpt_family="AT-rich"
                    15022..15327
repeat_region      /rpt_family="Alu"
                    15551..15617
repeat_region      /rpt_family="Alu"
                    16049..16068
misc_feature        /note="match to EST N71502 (NID:91228214) y*32a12.r1"
                    17185..17211
repeat_region      /rpt_family="(TTTTC)n"
                    17212..17495
repeat_region      /rpt_family="Alu"
                    18024..18155
repeat_region      /rpt_family="Alu"
                    18190..18236
repeat_region      /rpt_family="L2"
                    18277..18493
repeat_region      /rpt_family="L2"
                    18942..19025
repeat_region      /rpt_family="MTR"
                    19608..19920
repeat_region      /rpt_family="Alu"
                    20351..20475
repeat_region      /rpt_family="Alu"
                    20586..20820
misc_feature        /note="match to EST AA326909 (NID:91979280)"
                    21520..21689
repeat_region      /rpt_family="L2"
                    21690..21809
repeat_region      /rpt_family="Alu"
                    21810..22923
repeat_region      /rpt_family="L2"
                    23090..23229
repeat_region      /rpt_family="MTR"
                    23582..23680
repeat_region      /rpt_family="L1"
                    23681..23854
repeat_region      /rpt_family="Alu"
                    23855..24135
repeat_region      /rpt_family="Alu"
                    24136..24278
repeat_region      /rpt_family="Alu"
                    24279..24373
repeat_region      /rpt_family="L1"
                    24374..24688
repeat_region      /rpt_family="Alu"
```

```
repeat_region      24689..24888
                    /rpt_family="L1"
repeat_region      24889..25286
                    /rpt_family="L1"
repeat_region      25287..25867
                    /rpt_family="L1"
repeat_region      25819..25967
                    /rpt_family="L1"

Query Match      58.4% Score 22.2; DB 9; Length 219436;
Best Local Similarity 77.1%; Pred. No. 87;
Matches 27; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      4      tctagaggaattcgagtgacatactttggcgc 38
          |||||  |||||  |||||  |||||  |||||  ||
Db 121342  tctagaggaattcgagtgacatacttttcattgc 121376

RESULT 6
LOCUS      CDC2L1S01      1060 bp      DNA
DEFINITION Homo sapiens P1RSURE protein kinase (CDC2L1) gene, exon 1.
ACCESSION  AF080684
VERSION    AF080684.1 GI:3978425
KEYWORDS
SEGMENT
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE  1 (bases 1 to 1060)
            Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
            Ambros,P.F. and Kidd,V.J.
            Duplication of a genomic region containing the cdc2L1-2 and
            MRP21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2
            Genome Res. 8 (9), 929-939 (1998)
9842414
JOURNAL    MEDLINE
REFERENCE  2 (bases 1 to 1060)
            Gururajan,R., Lahli,J.L., Grenet,J., Easton,G., Gruber,I.,
            Ambros,P. and Kidd,V.J.
            Direct Submission
            Submitted (27-JUL-1998) Tumor Cell Biology, St Jude Children's
            Research Hospital, 332 N. Lauderdale St, Memphis, TN 38105, USA
FEATURES
            source
                1..1060
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="1"
                /map="1p36.3"
                <398..496
                /gene="CDC2L1"
                /number=1
BASE COUNT      205 a      299 c      341 g      215 t
ORIGIN
Query Match      57.9% Score 22; DB 9; Length 1060;
Best Local Similarity 100.0%; Pred. No. 11e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17      gaagtgcacatactttggcgc 38
          |||||  |||||  |||||  |||||  |||||  ||
Db      404      GAAGTGACGATCTTTGGCGC 425

RESULT 7
AF067513      2226 bp      mRNA
DEFINITION    Homo sapiens P1RSURE protein kinase alpha SV4 isoform (CDC2L1)
ACCESSION     AF067513
VERSION       AF067513.1 GI:3850305
KEYWORDS
```

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES

SOURCE

Location/Qualifiers
1. 41984
/organism="Caenorhabditis briggsae"
/strain="GuJARat G16"
/db_xref="taxon:6238"
/clone="G42E09"
ORIGIN

BASE COUNT 13435 a 7657 c 8328 g 12564 t

Query Match 60.0%; Score 22.8; DB 3; Length 41984;
Best Local Similarity 79.4%; Pred. No. 50;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 tgcctagaggaatcgaaatgacatactttg 34
||||| ||||| ||||| ||||| |||||
Db 3820 tgcctagaggaatcgaaatgacatactttg 3853

RESULT 5

AC005046 219436 bp DNA PRI 02-OCT-2000

LOCUS AC005046 Homo sapiens BAC clone CTB-13F3 from 7q422, complete sequence.

DEFINITION AC005046

VERSION AC005046.3 GI:6094632

KEYWORDS HTG.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 219436)

AUTHORS Sulston, J.E. and Waterston, R.

TITLE Toward a complete human genome sequence

JOURNAL 99063792

MEDLINE 2 (bases 1 to 219436)

REFERENCE Leonard, S. and Miller, N.

AUTHORS The sequence of Homo sapiens BAC clone CTB-13F3

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 219436)

AUTHORS Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington

MO 63108 USA University School of Medicine, 4444 Forest Park Parkway, St. Louis,

REFERENCE 4 (bases 1 to 219436)

AUTHORS Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (29-JUL-1999) Genome Sequencing Center, Washington

MO 63108 USA University School of Medicine, 4444 Forest Park Parkway, St. Louis,

REFERENCE 5 (bases 1 to 219436)

AUTHORS Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (22-OCT-1999) Genome Sequencing Center, Washington

MO 63108 USA University School of Medicine, 4444 Forest Park Parkway, St. Louis,

REFERENCE 6 (bases 1 to 219436)

AUTHORS Waterston, R.

TITLE Direct Submission

JOURNAL Submitted (02-OCT-2000) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Oct 22, 1999 this sequence version replaced gi:5649376.

COMMENT

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>

Contact: saplens@watson.wustl.edu
----- Summary Statistics
Center project name: H_RG013F03

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTB-13F3 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).

VECTOR: pBel0BAC11

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is CTA-364P16, 200 bp overlap. Actual start of this clone is at base position 83442 of CTA-364P16; actual end is at base position 219436 of CTB-13F3.

The clone CTB-13F3 contains three transposons which have been omitted from the submitted sequence. The transposons should insert after base position 14419, 66205, and 79419 respectively.

FEATURES

source

1. 219436
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q42"
/clone="CTB-13F3"
/clone_lib="CTB-978SK-B"
2179. .2254
/rpt_family="MIR"
2268. .2541
/rpt_family="Alu"
2948. .3174
/rpt_family="MIR"
3547. .3583
/rpt_family="AT_rich"
3705. .3730
/rpt_family="CAAAn"
4127. .4151
/rpt_family="AT_rich"
4154. .4271
/rpt_family="Alu"
4793. .4900
/rpt_family="MER1_type"
6540. .6582
/rpt_family="T-rich"
6583. .6899
repeat_region

----- Summary Statistics -----
Contig length: 184447
Phrap values in estimate: 183714
Average error rate (RCM-Phrap estimate): 0.000257953
Fraction of Phrap values less than 40 : 0.0455001
Number of consensus changing edits: 30
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
7134 catatattgc(n)taagtttttt catatattgc(t)taagtttttt
18131 gatcatatgc(n)aaaaaataaa gatcatatgc(a)aaaaaataaa
25646 cccacagctg(n)caacaacac cccacagctg(a)caacaacac
34492 tactaaatt(n)caaaaaaaa tactaaatt(a)caaaaaaaa
44313 actggtacc(n)ccatgacag actggtacc(c)ccatgacag
44690 ggcacaagac(n)ngggagacag ggcacaagac(t)ngggagacag
44691 gccaaaagac(n)gggagacagt gccaaaagac(g)gggagacagt
93098 ttaaaaaag(n)aacaagcata ttaaaaaag(a)aacaagcata
93106 gnaacaagc(n)taagacaag ttaaaaaagc(a)taagacaag
93155 tttaaaagc(n)ctaattcaaa tttaaaagc(a)ctaattcaaa
133602 tttttttt(g)taactttttt tttttttt(t)taactttttt
133605 actttttt(n)tttaactttt actttttt(l)tttaactttt
139683 ggtggtg(n)tcacacctgt ggtggtg(c)tcacacctgt
139946 gcaaaactcc(n)ctnaaaaaaa gcaaaactcc(a)ctnaaaaaaa
139949 aaactcnc(n)taaaaaataa aaactcnc(a)taaaaaataa
148974 gtagatgga(n)ntgtgacata gtagatgga(a)ntgtgacata
148975 tagatgtga(n)tgtagacata tagatgtga(a)tgtagacata
149511 tgaagtgaag(n)gtgtagaat tgaagtgaag(t)gtgtagaat
151700 ctccctggaac(n)cttaaggaat ctccctggaac(t)cttaaggaat
154316 tataccctta(n)nmctgtcct tataccctta(a)nmctgtcct
154317 ataccctta(n)nmctgtcct ataccctta(c)nmctgtcct
154318 taacctann(n)notgtcctc taacctann(c)notgtcctc
154319 acccttann(n)ctgtcctc acccttann(a)ctgtcctc
155088 ttgtgaatg(n)aaaaacatac ttgtgaatg(t)aaaaacatac
156338 cgttcagat(n)taaggaaatac cgttcagat(t)taaggaaatac
157074 taaaaaatlg(n)catgcaatc taaaaaatlg(t)catgcaatc
157110 ttgtcaacag(n)gctgccttt ttgtcaacag(t)gctgccttt
157188 acatagttta(n)aaacatag acatagttta(g)aaacatag
171922 aaatgactc(n)gaccttagag aaatgactc(t)gaccttagag
178161 catgttagcc(n)ggcaggtcct catgttagcc(a)ggcaggtcct

----- Distribution of Quality < 40 Bases -----
1000 | * * *
900 | * * *
800 | * * *
700 | * * *
600 | * * *
500 | * * *
400 | * * *
300 | * * *
200 | * * *
100 | * * *
0 | * * *
5 10 15 20 25 30 35 40
Phrap Value Range

----- FEATURES -----
Version: 1.01 gxf.
Location/Qualifiers
1. 184447
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP1-1036F1"
1. 68
repeat_region
/rpt_family="MER66B"

repeat_region complement(329, .389)
repeat_region /rpt_family="MIR"
repeat_region 499, .524
repeat_region /rpt_family="(CAAG)n"
repeat_region 1160, .1585
repeat_region /rpt_family="L1MC5"
repeat_region complement(1592, .1673)
repeat_region /rpt_family="AluSq"
repeat_region 1676, .1725
repeat_region /rpt_family="AluSx"
repeat_region 1727, .1760
repeat_region /rpt_family="(TCTA)n"
repeat_region
repeat_region /rpt_family="AT-rich"
repeat_region complement(2080, .2370)
repeat_region /rpt_family="AluY"
repeat_region complement(2395, .2679)
repeat_region /rpt_family="AluSx"
repeat_region 3700, .4548
repeat_region /rpt_family="L2"
repeat_region 4540, .4594
repeat_region /rpt_family="L2"
repeat_region 4604, .4624
repeat_region /rpt_family="(TCAA)n"
repeat_region 5423, .5931
repeat_region /rpt_family="L2"
repeat_region 6112, .6466
repeat_region /rpt_family="MLTII"
repeat_region complement(6525, .6638)
repeat_region /rpt_family="MIR"
repeat_region 6647, .6673
repeat_region /rpt_family="(TTG)n"
repeat_region complement(6674, .6963)
repeat_region /rpt_family="AluO"
repeat_region complement(6964, .7122)
repeat_region /rpt_family="MIR"
repeat_region complement(7138, .7440)
repeat_region /rpt_family="AluY"
repeat_region 7460, .7493
repeat_region /rpt_family="(A)n"

Query Match 61.6%; Score 23.4; DB 9; Length 184447;
Best Local Similarity 81.8%; Pred. No. 28;
Matches 27; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 3 cctagaggaattcgaagtacgacttctg 35
Db 74647 CTCCTAGAGCAATGCACTAACAACTGCTATG 74615

RESULT 4
CBRG42E09 41984 bp DNA INV 04-NOV-2000
LOCUS
DEFINITION Caenorhabditis briggsae cosmid G42E09, complete sequence.
ACCESSION AC084593
VERSION AC084593.1 GI:11095043
KEYWORDS HTG.
SOURCE
ORGANISM Caenorhabditis briggsae.
Caenorhabditis briggsae
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida;
Rhabdilitidae; Rhabdilitidae; Peloderinae; Caenorhabditis.
REFERENCE
AUTHORS Washington University Genome Sequencing Center.
TITLE The C. briggsae Genome Sequencing Project.
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 41984)
TITLE Direct Submission
JOURNAL Submitted (04-NOV-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
Submitted by: Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA
e-mail: jspiehl@watson.wustl.edu

CC	*	71777	71876:	gap of unknown length
CC	*	71877	84981:	contig of 13105 bp in length
CC	*	84982	85081:	gap of unknown length
CC	*	85082	97556:	contig of 124/5 bp in length
CC	*	97557	97656:	gap of unknown length
CC	*	97657	111043:	contig of 13387 bp in length
CC	*	111044	111143:	gap of unknown length
CC	*	111144	122275:	contig of 11132 bp in length
CC	*	122276	122375:	gap of unknown length
CC	*	123237	133263:	contig of 10888 bp in length
CC	*	133264	133363:	gap of unknown length
CC	*	133364	140049:	contig of 6686 bp in length
CC	*	140050	140149:	gap of unknown length
CC	*	140150	147482:	contig of 7333 bp in length
CC	*	147483	147582:	gap of unknown length
CC	*	147583	151952:	contig of 4370 bp in length
CC	*	151953	152052:	gap of unknown length
CC	*	152053	156921:	contig of 4865 bp in length
CC	*	156922	157021:	gap of unknown length
CC	*	157022	158420:	contig of 2399 bp in length
CC	*	158421	159520:	gap of unknown length
CC	*	159521	163612:	contig of 4093 bp in length
CC	*	163613	163712:	gap of unknown length
CC	*	163713	167217:	contig of 3505 bp in length
CC	*	167218	167317:	gap of unknown length
CC	*	167318	169954:	contig of 2637 bp in length
CC	*	169955	170054:	gap of unknown length
CC	*	170055	171664:	contig of 1610 bp in length
CC	*	171665	171764:	gap of unknown length
CC	*	171765	173324:	contig of 1560 bp in length
CC	*	173325	173424:	gap of unknown length
CC	*	173425	175056:	contig of 1632 bp in length
CC	*	175057	175156:	gap of unknown length
CC	*	175157	176359:	contig of 1203 bp in length.
XX				
Key			Location/Qualifiers	
FF				
FF	source	1..176359		
FF		/chromosome="3"		
FF		/db_xref="taxon:9606"		
FF		/organism="Homo sapiens"		
FF		/clone="RP11-360G1"		
XX				
Sequence		176359 BP; 53121 A; 34541 C; 34264 G; 52495 T; 1938 other.		

Query Match	61.6%	Score 23.4	DB 33	Length 176359
Best Local Similarity	81.8%	Pred. NO. 28		
Matches 27: Conservative	0	Mismatches	6	Indels 0
				Gaps 0
QY	3	ctctagagaatcgaagtgcagcactcttg	35	
Db	36576	CTCTAGAGGAATTGAGTAACATACGATG	36544	
RESULT 3				
AC012519/c				
LOCUS	AC012519	184447 bp	DNA	
DEFINITION	Homo sapiens 3 BAC RP11-1036F1 (Roswell Park Cancer Institute Human		PRI	31-AUG-2000
	BAC library) complete sequence.			
ACCESSION	AC012519			
VERSION	AC012519.9	GI:9558566		
KEYWORDS	HTG.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1 (bases 1 to 184447)			
AUTHORS	Muzny D.M., Adams C., Bailey M., Barbara J., Blankenburg K.,			
	Bodala B., Bouck J.J., Bowie S., Brooks A., Buhay C., Bunac C.,			
	Burkett C., Burrows J., Carter M., Chacko J., Chen Z., Cox C.,			
	David R., Delgado O., Desnazo D., Ding Y., Domah-Rashid N.,			
	Dugan-Rocha S., Durbin K.J., Fernandez C., Ferraguto D.,			

	Gorrell,J.H., Guevara,W., Harris,K., He,X., Hernandez,J., Hodgson,A., Hognes,M., Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M., Kelly,S., Kondyewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z., Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozano,R.J., Lu,J., Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Moorish,T., Moirish,T., Morgan,M., Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M., Relter,D., Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Succiang,R., Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wabnah,M., Wallington,S., Weinstock,G., Weinstein,I.R., Williamson,A., Worley,K., Wren,J., Wensford,G., Yu,W., Zhou,X., Naylor,S.L., Nelson,D.
TITLE	Unpublished Direct Submission
JOURNAL REFERENCE	2 (pages 1 to 184447)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL REFERENCE	Submitted (29-OCT-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AUTHORS	3 (pages 1 to 184447) Worley,K.C.
TITLE	Direct Submission
JOURNAL REFERENCE	Submitted (28-JUL-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA 4 (bases 1 to 184447)
AUTHORS	Worley K.C.
TITLE	Direct Submission
JOURNAL REFERENCE	Submitted (31-AUG-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Jul 28, 2000 this sequence version replaced gi:8694924.

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by Blast (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: this sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

TITLE Direct Submission
JOURNAL Submitted (10-AUG-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Aug 10, 2001 this sequence version replaced gi:8699801.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: <http://www.genome.washington.edu>
Contact: uwgchgs@u.washington.edu
Drafting Center: BCM
----- Project Information
Center project name: chr-3
Center clone name: RP11-360G1 (bc0384)
----- Summary Statistics
Sequencing vector: M13; 108821; 50% of reads
Sequencing vector: plasmid; 108752; 50% of reads
Chemistry: Dye-primer Bodipy; 2% of reads
Chemistry: Dye-terminator ET; 32% of reads
Chemistry: Dye-terminator Big Dye; 66% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 173927 bases at least Q40
Consensus quality: 174208 bases at least Q30
Consensus quality: 174306 bases at least Q20
Insert size: 174350; sum-of-ctrls
Quality coverage: 8.0x in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 6687: contig of 6687 bp in length
* 6688 6787: gap of unknown length
* 6788 174450: contig of 167663 bp in length.
* Location/Qualifiers
1..174450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-360G1"
/clone_lib="RP11 human BMC library 11"
1..6687
/note="assembly_name:Contig18"
misc_feature 6788..174450
/note="assembly_name:Contig19"
BASE COUNT 53199 a 34138 c 34305 g 52708 t 100 others
ORIGIN
Query Match 61.6%; Score 23.4; DB 2; Length 174450;
Best Local Similarity 81.8%; Pred. No. 28;
Matches 27; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 cttctaggaatcgaagcgacatctttgg 35
|||||
Db 75606 CTTCTAGGAAATGAAGTAACTACATACGTATGCG 75574

RESULT 2
AC055709/c
ID AC055709 standard; DNA: HTG; 176359 BP.
XX AC055709;
XX AC055709;
XX AC055709.7
XX
DT 20-APR-2000 (rel. 63 Created)
DT 07-SEP-2000 (rel. 65; last updated, Version 11)
XX
XX Homo sapiens chromosome 3 clone RP11-360G1, WORKING DRAFT SEQUENCE, 20
DE

DE unordered pieces.
XX HTG; HTGS_DRAFT; HTGS_PHASE1.
KW
XX
OC Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
XX
XX
XX [1]
XX 1-176359
RA Muzny D.M., Adams C., Bailey M., Barabara J., Blankenburg K., Bodda B.,
RA Bouck J., Boyle S., Brooks A., Buhay C., Bunac C., Burkett C., Burrows J.,
RA Carter M., Chacko J., Chen Z., Cox C., David R., Delgado O., Deshazo D.,
RA Ding Y., Domah-Rashid N., Dugan-Rocha S., Durin K.J., Fernandez C.,
RA Ferraguto D., Forcum-Ranney J., Frantz P., Ganesh R., Gorrell J.H.,
RA Gorrell L.L., Guevara W., Harris K., Hernandez J., Hodgson A., Hognes M.,
RA Holloway C., Hosack H., Jackson L.E., Jackson L., Jia Y., Jones M.,
RA Kelly S., Kondejewski N., Kong Y., Kovar C., Leal B., Li Z., Lichtarge O.,
RA Liu J., Liu W., Logan O., Lozado R.J., Lu J., Lucier R., Martin R.,
RA Martinez C., McLeod M.P., Mel G., Morgan M., Morris S., Nash S., Nelson A.,
RA Nguyen R., Nguyen N., Nguyen S., Oswald G., Parish B., Paxton S., Payton B.,
RA Perez L., Pu L.L., Quiles M., Reiter D., Rives M., Samuel S., Say J.,
RA Scherer S., Shah E., Shen H., Simon M., Sparks A., Stamps A., Sugang R.,
RA Tabor P., Taylor T., Vasquez L., Vinson R., Vo Q., Wabash M.,
RA Watlington S., Weinstock G., Weinstock I.R., Williamson A., Worley K.,
RA Wren J., Wrensford G., Yu W., Zhou X., Nelson D., Gilbs R.;
RT "Direct Submission";
XX unpublished.
XX [2]
XX 1-176359
XX Worley K.C.;
RT
RT Submitted (18-APR-2000) to the EMBL/GenBank/DBJ databases.
RL Human Genome Sequencing Center, Department of Molecular and Human Genetics,
RL Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
XX
XX
XX On Jun 25, 2000 this sequence version replaced gi:8102160.
CC ----- Genome Center
CC Center: Baylor College of Medicine
CC Center code: BCM
CC Web site: <http://www.hgsc.bcm.tmc.edu/>
CC Contact: hgsc-help@bcm.tmc.edu
CC ----- Project Information
CC Center project name: HAP0
CC Center clone name: RP11-360G1
CC ----- Summary Statistics
CC Sequencing vector: M13; 108821
CC Chemistry: Dye-primer Bodipy; 4% of reads
CC Chemistry: Dye-terminator Big Dye; 96% of reads
CC Assembly program: Phrap; version 0.990329
CC Consensus quality: 156972 bases at least Q40
CC Consensus quality: 165022 bases at least Q30
CC Consensus quality: 169145 bases at least Q20
CC Estimated insert size: 170012; sum-of-ctrls estimation
CC Quality coverage: 0x in Q20 bases; agarose-TP estimation
CC Quality coverage: 3.9x in Q20 bases; sum-of-ctrls estimation
CC -----
CC * NOTE: Estimated insert size may differ from sequence length
CC (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
CC * NOTE: This is a 'working draft' sequence. It currently
CC consists of 20 contigs. The true order of the pieces
CC is not known and their order in this sequence record is
CC arbitrary. Gaps between the contigs are represented as
CC runs of N, but the exact sizes of the gaps are unknown.
CC * This record will be updated with the finished sequence
CC as soon as it is available and the accession number will
CC be preserved.
CC 1 31369: contig of 31369 bp in length
CC * 31370 31469: gap of unknown length
CC * 31470 54720: contig of 23251 bp in length
CC * 54721 54820: gap of unknown length
CC * 54821 71776: contig of 16956 bp in length
CC

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:54:48 ; Search time 2399.05 Seconds
(without alignments)
261.309 Million cell updates/sec

Title: US-09-915-060-7

Perfect score: 38
Sequence: 1 tgccttagaggaattcgaagtacagatactttggcgc 38

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pt: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vi: *
30: em_htgo_hum: *
31: em_htgo_inh: *
32: em_htgo_rod: *
33: em_htg_hum: *
34: em_htg_inh: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	23.4	61.6	174450	2	AC093115	AC093115 Homo sapi
C 2	23.4	61.6	176359	33	AC055709	AC055709 Homo sapi
C 3	23.4	61.6	184447	9	AC012519	AC012519 Homo sapi
C 4	22.8	60.0	41984	3	CBRG42E09	AC084593 Caenorhab
5	22.2	58.4	219436	6	AC005046	AC005046 Homo sapi
6	22	57.9	1060	9	CD22L1S01	AF080684 Homo sapi
7	22	57.9	2226	9	AF067513	AF067513 Homo sapi
8	22	57.9	2340	9	HSU04818	U04818 Human prote
9	22	57.9	2349	9	AF067517	AF067517 Homo sapi
10	22	57.9	2444	9	HSU04817	U04817 Human prote
11	22	57.9	2465	9	HSU07705	U07705 Human prote
12	22	57.9	2471	6	AX033422	AX033422 Sequence
13	22	57.9	2471	9	HSU04816	U04816 Human prote
14	22	57.9	2471	9	HSU07704	U07704 Human prote
15	22	57.9	2477	9	HSU04824	U04824 Human prote
16	22	57.9	2486	9	AF067512	AF067512 Homo sapi
17	22	57.9	2500	9	AF067514	AF067514 Homo sapi
18	22	57.9	2525	9	AF067515	AF067515 Homo sapi
19	22	57.9	2533	9	AF067516	AF067516 Homo sapi
20	21.8	57.4	18061	9	HSCYTOK20	X73501 H. sapiens g
C 21	21.8	57.4	137022	9	AC004231	AC004231 Homo sapi
C 22	21.6	56.8	8171	3	DMVIN	X96601 D.melanogas
C 23	21.6	56.8	17254	3	DROPCXGEN	M74329 Drosophilla
C 24	21.6	56.8	26697	3	DMC103B4	AL009193 Drosophill
C 25	21.6	56.8	36824	8	AP000372	AP000372 Arabidops
C 26	21.6	56.8	38892	2	DMC62D9	AL009171 Drosophill
C 27	21.6	56.8	46221	2	AC017684	AL017684 Drosophill
28	21.6	56.8	116524	9	AL138686	AL138686 Human DNA
29	21.6	56.8	175188	2	AL445238	AL445238 Homo sapi
C 30	21.6	56.8	291976	3	AE003423	AE003423 Drosophill
C 31	21.4	56.3	37049	1	MLNCEA	Z46257 M.leptae ac
C 32	21.4	56.3	37711	3	CEZK896	Z82288 Caenorhabdi
C 33	21.4	56.3	86825	9	AC006967	AC006967 Homo sapi
34	21.4	56.3	342300	1	MLEPRTN8	AL583924 Mycobacte
C 35	21.2	55.8	652	3	AF176697	AF176697 Circulife
C 36	21.2	55.8	66224	9	AC073884	AC073884 Homo sapi
C 37	21.2	55.8	159681	9	AC025160	AC025160 Homo sapi
C 38	21.2	55.8	172893	2	AC024586	AC024586 Homo sapi
C 39	21	55.3	36783	1	AF078135	AF078135 Leptospi
C 40	21	55.3	46057	9	AC073276	AC073276 Homo sapi
C 41	21	55.3	145603	9	AL157768	AL157768 Human DNA
C 42	21	55.3	169581	2	AC007490	AC007490 Homo sapi
C 43	21	55.3	193539	2	AC022190	AC022190 Homo sapi
C 44	20.8	54.7	2945	3	D83384	D83384 Tetrahymena
45	20.8	54.7	5108	1	BB044914	U44914 Borrelia bu

ALIGNMENTS

RESULT 1	AC093115	174450 bp	DNA	HTG	10-AUG-2001
AC093115/C	AC093115	174450 bp	DNA	HTG	10-AUG-2001
LOCUS	AC093115	174450 bp	DNA	HTG	10-AUG-2001
DEFINITION	AC093115	174450 bp	DNA	HTG	10-AUG-2001
ACCESSION	AC093115	174450 bp	DNA	HTG	10-AUG-2001
VERSION	AC093115.1	174450 bp	DNA	HTG	10-AUG-2001
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
	Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and				
	Haugen,E.D.				
TITLE	Unpublished				
JOURNAL	2 (bases 1 to 174450)				
REFERENCE	Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and				
AUTHORS	Haugen,E.D.				

us-09-915-060-1.rge

Page 9

D_b 1027 GAAATATTACAGAGGAGAAGCAAGAGAGAGAGAGACCGGCAGCAACTCTTGAGSAG 1086
Q_y 181 gcatcagagcaactctcgccgaagaagtaagtgtgagaaatg 222
D_b 1087 GCATCAGAGCAGTCTGCCGAAGAAGTAAGTAGAAGAAAGATG 1128

```
Search completed: December 4, 2001, 18:54:30
Job time: 9166 sec
```


Db 1000 GAAGAAATCAGAGCAGAGGAAGAGAGAGAGAGAGAGACCGCAGCACTCTGAGAG 1059

QY 181 gcatcagagcagctgccgaagaagtaagtgaagaagaatg 222
|||||

Db 1060 GCATCAGAGCAGTCTGCCGAGAGTAAGTGAGAGAAATG 1101

RESULT 12
AF067522 2448 bp mRNA PRI 07-NOV-1998
LOCUS AF067522
DEFINITION Homo sapiens PITSURE protein kinase beta SV6 isoform (CDC2L2) mRNA,
complete cds.
AF067522
VERSION AF067522.1 GI:3850323
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2448)
AUTHORS Gururajan,R., Laheti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2
JOURNAL Genome Res. 8 (9), 929-939 (1998)
MEDLINE 98424414
REFERENCE 2 (bases 1 to 2448)
AUTHORS Gururajan,R., Laheti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P., and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
MEDLINE Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
94253170
FEATURES
source location/Qualifiers
1..2448
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2448
/gene="CDC2L2"
80..2422
/gene="CDC2L2"
/note="corresponding genomic sequence deposited as GenBank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PITSURE protein kinase beta SV6 isoform"
/protein_id="AAC72087.1"
/translation="MGDEKDSWKVKTLDLILQEKRRRQEKAEIKRLKNSDRDSC
RSLSEGLDHCMEITIRNSPYRSDMSDRGEDDSLAIKPPOQMSREKYNHRKD
EKREKKHARVKEHREHRRKRRREODKARREMRQKREMARSHRSRRDRDLQLE
RKREERKMRQEKQEQREKREKRRRAREKREARREVSANHRTREDSYSDYKASH
WSRSPPRRPRERELQDGRPGEARPARAKPAQLKEEKEERDLSDLODISDSERK
TSSAESSSAEGSESEEBEEDSSTSESEEBEEDSSTSESEEBEEDSSTSESEEBE
EYSE
POSPALPLTEIKOELPKYLPALOGCRSVDFEFCINRIEGTGYVYRAKDKTDEIVA
LKLTKMEKEKEGSPITSIREINTILKAQHPNIVTVREIYVGSNMCKIYIVMNVENHL
KSLMTWKQEPFLGPEVKTLMIDLGRVKHLHNMILHRLKSNLISAGLTKGDEGLAE
GLAREYSPKLKATPVVVTQWYRAPLGLGAKETSTAVDMWSVGCIFGELTLQKPLP
GNSIEDIINKVFEKLTSPSEKIMPGYSELVYVKMTFSEHPNNLRKRFGLLSDDGF
DLANKFLTYPPGRRISAEGLKHEYFRETPLPIDSMFTPMWPAKSEQQVYKRGTSPPR
PEGGLGSGDGDILKETGFHLTTTNGASAGPGFSLK"

BASE COUNT 694 a 598 c 796 g 360 t

ORIGIN

Query Match 100.0%; Score 222; DB 9; Length 2448;
Best local Similarity 100.0%; Pred. No. 1,1e-35;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 60
|||||

Db 884 GACATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 943

QY 61 tcaggtctgaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 120
|||||

Db 944 TCAGGTTCTGAG 1003

QY 121 gaagaatcagagcagctgccgaagaagtaagtgaagaagaatg 222
|||||

Db 1004 GAAGAAATCAGAGCAGAGGAAGAGAGAGAGAGAGAGACCGCAGCACTCTGAGAG 1063

QY 181 gcatcagagcagctgccgaagaagtaagtgaagaagaatg 222
|||||

Db 1064 GCATCAGAGCAGTCTGCCGAGAGTAAGTGAGAGAAATG 1105

RESULT 13
HSU07705 2465 bp mRNA PRI 08-JUL-1994
LOCUS HSU07705
DEFINITION Human protein kinase PITSURE isoform PBETA22 (p58CDC2L1) mRNA,
complete cds.
U07705
VERSION U07705.1 GI:507428
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2465)
AUTHORS Xiang,J., Laheti,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE Molecular cloning and expression of alternatively spliced PITSURE
protein kinase isoforms
JOURNAL J. Biol. Chem. 269, 15786-15794 (1994)
MEDLINE 94253170
REFERENCE 2 (bases 1 to 2465)
AUTHORS Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-1994) Vincent J. Kidd, St. Jude Children's
MEDLINE Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
94253170 Memphis, TN 38101, USA
FEATURES
source location/Qualifiers
1..2465
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36-2"
/cell_line="Hela S3"
/cell_type="epitheloid"
/tissue_type="cervix"
/clone_id="Hela cDNA library"
/dev_stage="adult"
1..2465
112..2439
/gene="p58CDC2L1"
112..2439
/gene="p58CDC2L1"
/note="p58CDC2L1"
/function="protein kinase"
/codon_start=1
/product="PITSURE isoform PBETA22"
/protein_id="AAI19595.1"
/translation="MGDEKDSWKVKTLDLILQEKRRRQEKAEIKRLKNSDRDSC
RSLSEGLDHCMEITIRNSPYRSDMSDRGEDDSLAIKPPOQMSREKYNHRKD
EKREKKHARVKEHREHRRKRRREODKARREMRQKREMARSHRSRRDRDLQLE
RKREERKMRQEKQEQREKREKRRRAREKREARREVSANHRTREDSYSDYKASH
WSRSPPRRPRERELQDGRPGEARPARAKPAQLKEEKEERDLSDLODISDSERK
TSSAESSSAEGSESEEBEEDSSTSESEEBEEDSSTSESEEBEEDSSTSESEEBE
EYSE
POSPALPLTEIKOELPKYLPALOGCRSVDFEFCINRIEGTGYVYRAKDKTDEIVA
LKLTKMEKEKEGSPITSIREINTILKAQHPNIVTVREIYVGSNMCKIYIVMNVENHL
KSLMTWKQEPFLGPEVKTLMIDLGRVKHLHNMILHRLKSNLISAGLTKGDEGLAE
YGSPLKATPVVVTQWYRAPLGLGAKETSTAVDMWSVGCIFGELTLQKPLP
GNSIEDIINKVFEKLTSPSEKIMPGYSELVYVKMTFSEHPNNLRKRFGLLSDDGF
DLANKFLTYPPGRRISAEGLKHEYFRETPLPIDSMFTPMWPAKSEQQVYKRGTSPPR
PEGGLGSGDGDILKETGFHLTTTNGASAGPGFSLK"

ACCESSION	AF067519	
VERSION	AF067519.1	GI:3850317
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 2439)	
AUTHORS	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P.F. and Kidd,V.J.	
TITLE	Duplication of a genomic region containing the Cdc2l1-2 and MPM21-12 genes on human chromosome 1p36.3 and their linkage to D12Z2	
JOURNAL	Genome Res.	8 (9), 929-939 (1998)
MEDLINE	98424414	
REFERENCE	2 (bases 1 to 2439)	
AUTHORS	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J.	
TITLE	Direct Submission	
JOURNAL	Submitted (18-MAY-1998)	
FEATURES	Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA	
source	Location/Qualifiers	
	1..2439	

```

gene
CD5
1. .2439
/gene="CDC2L2"
80. .2413
/gene="CDC2L2"
/note="corresponding genomic sequence deposited as GenBank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PITSLRE protein kinase beta SV1 isoform"
/protein_id="A072084.1"
/db_xref="GI:3850318"
/translation="MDGEKSWYKLTDEILQEKRRKEOEKAELIKLNSDDRSK
RDLSEGLDHCHMEITIRNSPYRREDSMDRGCEDSLAIKPKQOQSWKEKYHAKRD
EKREKRRHHSHSAGEGKHARKERERERKRHHREEDQKRRREERQKRRREMAHESR
REKRLLEOLERKERREKRMKROOEKROEKREERERARERERAREREVAHNRHTESE
DYSDQVKAHMSWRSPPRRPRERPELTGCRPKVPEKKMEEDLSDLDIDSEKRTSS
AESSAASGSGSEBEEBEEBEGTSTSESEBEEBEEBEEBEEBEEBEEBEEBEEBEEB
SEEMSEDERENENHLLVVPESRFDKSGSESEAEVEEGTPOSSALTEGDIYAPDS
PALPLEIKPELPRALQOGRSVEFQINRIIEGTGVVYRAKDKIDEIYALKR
LKNKEKEGFPITSLREINTILKQAQHFOITVREIIVGSSNMKTYIYMYNVEDHLSL
MEMKQOPLFGEVETLMIOLLRGVHLHDWMLIHLRDLKTSNLLISHAGILIKVDFEGLA
MEMKQPLKATTPVYVQWRAPELLGAKRTSYAVDMWSGCIJGELLTPQKPLPFGMS
EIDQINKVEKEITPSEKIMPGYSLELPPVYKMTFSEHPYNNLKKRFALLSDQGFIDLMI
NKLITLPPGRIRISAEOLKHEITRETLPLDPSMPLPVPKASEQOVRKGTSPRPPEG
GLDLSQGDNDLKEITGHFTLTNQGASAAAPGFSLP"
BASE COUNT
695 a 592 c 790 g 362 t
BRIGIN

```

[illegible]

RESULT	11			
LOCUS	HSU04817	2444 bp	mRNA	PRI 08-JUL-1994
DEFINITION	Human protein kinase PITSLRE alpha 2-3 mRNA, complete cds.			
ACCESSION	U04817			
VERSION	U04817.1	GI:507161		
KEYWORDS	.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 2444)			
TITLE	Xiang,J., Lahli,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J. Molecular cloning and expression of alternatively spliced PITSLRE protein kinase isoforms			
JOURNAL	J. Biol. Chem. 269, 15786-15794 (1994)			
MEDLINE	94253170			
REFERENCE	2 (bases 1 to 2444)			
AUTHORS	Kidd,V.J.			
TITLE	Direct Submission			
JOURNAL	Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St., Memphis, TN 38101, USA			
FEATURES	Location/Qualifiers			
source	1..2444			

```

mRNA
CDS
1..2448
112..2418
/codon_start=1
/product="pTSLR alpha 2-3"
/protein_id="AA19583.1"
/db_xref="GI:507162"
/translation="MGEKDSWKVLTDEILLOEKKRRKEOEKAETIKLNSDROSK
RDSIEGELDHCMETIRNSPYAREDSMDRGEEDSLAIKPPQOASREKRYHNRD
EKREKKHARVKKKEHEHRRKRRREEDQARKEWERQKREMAREHREPRLEOL
EKRSERKPRKEOEQEOREQREKREKREBERK EKEAREVYSAHRTREYSDKVAS
HWSSRPPRRPERFEILDGRRKPVYKEEMERDLSDLODLSHRTREYSAESSAEG
SGSEEEHEEEEEEESESESESESEEESEEESESESESESESESESESESESESE
EREHEHLLVVPESRFDPSGSESEEAEEVEEGEPQSSALTEGQYVDVDPALSPLEK
QELPKYLPALOGSRVSEPOCINRIEGTGGVYARADCKRTDEIVALAKIKMEKRG
PPLTSEPLNTILKAQHPNTPVAREIYGSNMOKYIVMNYVEIDLKSLMETMKOPL
PGVYKTLMDILGSKVRLHNDNMILHRDKLSNLLSAGILKLGDEFLAEYSGPLKA
YTPVYVTLMTAPBELGAKETSTAVDMNMGVGLTFGLLQKLPFEEKSEIDQIKVFA
KIDGTSEKIMWPGISLPAVKMTFSPVNNYLKRGFALLSDQGDLMNKPFLTPYFG
RRIISAEGLDHEYFRETPIDPSMFTWPAKSEQKVVYKRGTSRPREGGGLGYSLQED
DLKEETGEHLTTTNOGSAAGPFSLKF"
596 c 798 g 357 t

```

Query Match	100.0%;	Score 222;	DB 9;	length 2444;
Best Local Similarity	100.0%;	Score No. 1, 1e-35;		
Matches	222;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

QY	1	gacatcagcgacagcgaggaagaccagctcgccgagtcctctgctacagcaatccagc	60
Db	880	GACATCAGCGCGACACGAGGAGGACACCGCTCGGCCGCGAGTCTCTGTAGCGAATTCAGGC	939
QY	61	tcaagctctgaggaaggaagagagaggaaggaaggaaggaaggaaggaaggaagcagc	120
Db	940	TCAGCTTCTGAGGAGAAAGAGAGAGAGAGGAAAGAGAGAGAGGAGGAGGAGGAGGAGGAGT	999
QY	121	gaagaaatcagaggaaggaaggaaggaaggaaggaaggaagcagcagcaactctgaagag	180

MEDLINE REFERENCE	(bases 1 to 2362)
98424414	

JOURNAL Genome F
MEDLINE 98424414

FEATURES		Location/Qualifiers
source		1..2323 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="1" /map="1p36.3" 1..2323 /gene="CDC2L2" 1104..2297 /gene="CDC2L2" /note="corresponding genomic sequence deposited as GenBank Accession Numbers AF080689-AF080697 and AF092426-AF092428"
gene		
CDS		
BASE COUNT	673 a 550 c 754 g 346 t	
ORIGIN	FSLK"	
Query Match	100.0%; Score 222; DB 9; Length 2323;	
Best Local Similarity	100.0%; Pred. No. 1,1e-35;	
Matches 222; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Oy	1 gacatcagcgaagaagaagaagaccagctcgcgcgagctcctcgttcagcagaatcagc 60 	
Db	875 GACATCAGCAGAGAGGAGAGAGAACCAAGCTCGGCCGAGTCTCTGACGAATCAGC 934 	
Oy	61 tcaagttctgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 120 	
Db	935 TCAGGTTCTGAG 994 	
Oy	121 gaagaatcagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 180 	
Db	995 GAAAGATCAG 1054 	
Oy	181 gcatcagaagcagctctgcgaagaagtaagtgaagaagaatg 222 	
Db	1055 GCATCAGAGCACTCTGCCAGAGAGAGTAACTGAGAGAAAGAT 1096 	
RESULT 5		
AF067525	2329 bp mRNA PRI 07-NOV-1998	
LOCUS	Homo sapiens PITSLRE protein kinase beta SV13 isoform (CDC2L2)	
DEFINITION	mRNA, complete cds.	
ACCESSION	AF067525	
VERSION	AF067525.1 GI:3850329	
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
AUTHORS	1 (bases 1 to 2329) Gururajan, R., Lahti, J.M., Grenet, J., Easton, J., Gruber, I., Ambros, P.F. and Kidd, V.J.	
TITLE	Duplication of a genomic region containing the Cdc2L1-2 and MPM21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2 genome Res. 8 (9), 929-939 (1998)	
JOURNAL	98424414	
MEDLINE	2 (bases 1 to 2329)	
REFERENCE	Gururajan, R., Lahti, J.M., Grenet, J., Easton, J., Gruber, I., Ambros, P. and Kidd, V.J.	
AUTHORS		
TITLE	Direct Submission	
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's	

Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA

FEATURES
source location/Qualifiers
1..2329 /organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2329 /gene="CDC2L2"
68..1654 /gene="CDC2L2"
/note="corresponding genomic sequence deposited as GenBank Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PITSLRE protein kinase beta SV13 isoform"
/protein_id="AAC72090.1"
/db_xref="GI:3850330"
/translation="METTIRSPYRRDSDMEDGEEEDSLAIRPOOMSKKEYHHNRKDEKRKKHARVAREHERERKRRKEODKARRREWQKREMARERSRRCRDGCLEFRDLPLERKERRERKRRMRQOEQRQKERERAERERAREVSANHTKRELDFSRKVASHWRSRPPPPRPFELADGRKPKVEEMERDLSDIDISSEKTSSAESASGSSEEEEEEEEEEESTSEESSEEEEEEETGSNSAEASQSASEVSESMSEDEREENNHLLVPDSRFSDSGSEEAEEGGTPSOALTEGVDPSPSPALLPIELKOELPKYLPAIOGCCRVPEEFCNRIEEGTGYVARAKTKTDIYALKRLMKEREKEGFPTTSILREINTILKAQHNYIVRYEIVYGSNMDKIYYIMNVYEHDKLSLETWKOPFLPGEVKTLMIQLRGVKHLDMNVLHRDLKTSMLLSHGICLVKPSPPSGPSODPPEGPHSRSPSYAG"

BASE COUNT 634 a 605 c 738 g 352 t

ORIGIN

Query Match 100.0%; Score 222; DB 9; Length 2329;
Best Local Similarity 100.0%; Pred. No. 1,Je-35;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dy 1 gacatcagcgaaagcgagaaggaagcaaccgctgcgccgagtcctcgtcacgaatccaggc 60
|||||
Db 689 GACATCAGCGAAGCAGGAGAAGCAAGCACACTGCGCGAGTCCTCTCATGACAAATCAGGC 748
|||||
Dy 61 tcaggtcttgaggaagagagagagaggaagagagagagagagagagagagagagagagc 120
|||||
Db 749 TCAGGTCTTGAGGAAGAGAAAGCAGGAGGAGGAAGAGGAGAGAGAGAGAGAGAGAGC 808
|||||
Dy 121 gaagaatcag 180
|||||
Db 809 GAAGATCAGACGAG 868
|||||
Dy 181 gatcagagcagctctgcgagaagatgaagtgaaggaagagagagagagagagagagagag 222
|||||
Db 869 GCATCAGACGACTGCGCAGAGAGAAAGTAAGTAGGAAGAAATG 910
|||||

RESULT 6

LOCUS HSU04818 2340 bp mRNA -08-JUL-1994
DEFINITION Human protein kinase PITSLRE alpha 2-4 mRNA, complete cds.
ACCESSION U04818
VERSION U04818.1 GI:507163
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2340)
AUTHORS Xiang,J., Iahji,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE Molecular cloning and expression of alternatively spliced PITSLRE protein kinase isoforms
JOURNAL J. Biol. Chem. 269, 15786-15794 (1994)
MEDLINE 94253170
REFERENCE 2 (bases 1 to 2340)
AUTHORS Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

40

Result	Query	Score	Match	Length	DB	ID	Description
1	222	100.0	222	6	AX033420	AX033420	Sequence
2	222	100.0	222	6	AX033421	AX033421	Sequence
3	222	100.0	660	6	AX033425	AX033425	Sequence
4	222	100.0	2333	9	AF067518	AF067518	Homo sapi
5	222	100.0	2330	9	AF067525	AF067525	Homo sapi
6	222	100.0	2249	9	HS004818	U04818	Human prote
7	222	100.0	2362	9	AF067523	AF067523	Homo sapi
8	222	100.0	2436	9	AF067521	AF067521	Homo sapi
9	222	100.0	2469	9	AF067520	AF067520	Homo sapi
10	222	100.0	2439	9	AF067519	U04817	Human prote
11	222	100.0	2449	9	HS004817	AF067522	Homo sapi
12	222	100.0	2448	9	AF067522	U07705	Human prote
13	222	100.0	2465	9	HS007705	AX033422	Sequence
14	222	100.0	2471	6	AX033422	U04816	Human prote
15	222	100.0	2471	9	HS004816	U07704	Human prote
16	222	100.0	2471	9	HS007704	U04824	Human prote
17	222	100.0	2477	9	HS004824	AF114497	Homo sapi
18	220.4	99.3	1275	9	AF114497	AF067513	Homo sapi
19	203	91.4	2226	9	AF067513	AF067517	Homo sapi
20	203	91.4	2349	9	AF067512	AF067512	Homo sapi
21	203	91.4	2486	9	AF067512	AF067514	Homo sapi
22	203	91.4	2500	9	AF067514	AF067515	Homo sapi
23	203	91.4	2525	9	AF067515	AF067516	Homo sapi
24	203	91.4	2533	9	AF067516	AK000081	Homo sapi
25	201.4	90.7	2544	9	AK000081	AF080669	Homo sapi
26	155.4	70.0	592	9	CDC2L2S08	AL031282	Human DNA
27	134.8	60.7	110608	9	HS283353	AF080678	Homo sapi
28	134.4	60.5	480	9	CDC3L1S08	M58693	Mouse p58/
29	128.6	57.9	1936	10	MUSP58BTA	L37092	Mouse muscul
30	127.8	57.6	3161	10	MUSCDBK	AX033426	Sequence
31	87	39.2	87	6	AX033426	AC087419	Mus muscu
32	76	34.2	40386	2	AC087419	BC006628	Mus muscu
33	75.4	34.0	2090	10	MCMEEN	X55038	Mouse MCEN
34	75.4	34.0	3840	10	MCMEEN	AC004914	Homo sapi
35	74.8	33.7	143769	9	AC004914	AC009601	Leishmani
36	72.8	32.8	39987	9	AC009601	AC005804	leishmani
37	72.8	32.8	42160	3	AC005804	AC040982	Mus muscu
38	72.8	32.8	215963	2	AC040982	AL583824	Homo sapi
39	72.4	32.6	161317	2	AL583824	Y13301	Homo sapien
40	72.2	32.5	13095	2	HSFGFR4G	AL591805	Mus muscu
41	71.8	32.3	243075	2	AL591805	AC025669	Mus muscu
42	71.6	32.3	203674	2	AC025669	AC091162	Homo sapi
43	71.2	32.1	149557	2	AC091162	AC073490	Homo sapi
44	71.2	32.1	150803	2	AC073490	AC083818	Mus muscu
45	71.2	32.1	209119	2	AC083818		

RESULT	1
AX033420	
LOCUS	
DEFINITION	AX033420 222 bp DNA PAT 21-SEP-2000
ACCESSION	Sequence 1 from Patent WO0044896.
VERSION	AX033420
KEYWORDS	AX033420.1 GI:10280181
SOURCE	.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (bases 1 to 222)
VIAAMS	Beyaert R. and Cornelis S.
INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORN	Internal ribosome entry site (IRES), vector containing same and uses thereof
SIGRID (BE)	Patient: WO 0044896-A 1 03-AUG-2000;
LOCATION/Qualifiers	VIAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORN
FEATURES	
SOURCE	1..222

CORPORATE SOURCE: Y. Furuchi, AGENE Research Institute, 200 Kajiwara, Kamakura, Kanagawa 247, Japan. furuchi@gene.co.jp
 SOURCE: Molecular and Cellular Biology, (1998) 18/11 (6191-6200).
 Refs: 49
 ISSN: 0270-7306 CODEN: MCEBD4
 United States
 DOCUMENT TYPE: Journal; Article
 FILE SEGMENT: 029 Clinical Biochemistry
 LANGUAGE: English
 SUMMARY LANGUAGE: English

L6 ANSWER 4 OF 4 BIOSIS COPYRIGHT 2001 BIOSIS
 TI **p53** Regulates the minimal promoter of the human topoisomerase II- α gene.
 AB DNA topoisomerase II- α is an essential enzyme for chromosome segregation during mitosis. Consistent with a cell division-specific role, the expression of the topoisomerase II- α gene is strongly influenced by the proliferation status of cells. The **p53** protein is one of the most important regulators of **cell cycle** progression in mammals, with an apparent dual role in the induction of **cell cycle** arrest following cytotoxic insults and in the regulation of the apoptotic cell death pathway. We have analysed whether **p53** plays a role in regulating expression of the human topoisomerase II- α gene. We show that wild-type, but not mutant, **p53** is able to decrease substantially the activity of the full length topoisomerase II- α gene promoter. Using a series of constructs comprising various deleted or mutated versions of the promoter lacking critical **cis-acting elements**, we show that this **p53**-specific regulation of the topoisomerase II- α promoter is independent of all characterized transcription factor binding sites and is directed at the minimal gene promoter. We conclude that expression of wild-type **p53** induces downregulation of the human topoisomerase II- α promoter by acting on the basal transcription machinery. These findings implicate topoisomerase II as one of the downstream targets for **p53**-dependent regulation of **cell cycle** progression in human cells.

ACCESSION NUMBER: 1997:19679 BIOSIS
 DOCUMENT NUMBER: PREV199799318882
 TITLE: **p53** Regulates the minimal promoter of the human topoisomerase II- α gene.
 AUTHOR(S): Sandri, M. Ines; Isaacs, Richard J.; Ongkeko, Weg M.; Harris, Adrian L.; Hickson, Ian D. (1); Brogini, Massimo; Vikhanskaya, Faina
 CORPORATE SOURCE: (1) Imperial Cancer Research Fund, Univ. Oxford, Inst. Molecular Med., John Radcliffe Hosp., Oxford OX3 9DU UK
 SOURCE: Nucleic Acids Research, (1996) Vol. 24, No. 22, pp. 4464-4470.
 ISSN: 0305-1048.
 Article
 LANGUAGE: English

=>

repression of the cdc2 promoter, while E1B-19K does not. Since the cdc2 promoter does not contain a TATA sequence, we performed deletion and point mutation analyses and identified the inverted CCAAT sequence located at -76 as a **cis-acting element** for the **p53**-mediated regulation. We found that a specific DNA-protein complex is formed at the CCAAT sequence and that this complex contains the NF-Y transcription factor. Consistently, a dominant negative mutant of the NF-YA subunit, NF-YA₂₉, decreases the cdc2 promoter, and **p53** does not further decrease the promoter activity in the presence of NF-YA₂₉. These results suggest that **p53** negatively regulates cdc2 transcription and that the NF-Y transcription factor is required for the **p53**-mediated regulation.

ACCESSION NUMBER: 2000:2169 BIOSIS
DOCUMENT NUMBER: PREV20000002169

TITLE: **p53** negatively regulates cdc2 transcription via

AUTHOR(S): Yun, Jeaho; Chae, Hee-Don; Choy, Hyon E.; Chung, Jongkyeong; Yoo, Hyang-Sook; Han, Moon-Hi; Shin, Deug Y.

CORPORATE SOURCE: (1) Bioscience Research Division, Korea Research Institute of Bioscience and Biotechnology, Yuseong, Taejeon, 305-600

SOURCE: South Korea

Journal of Biological Chemistry, (Oct. 15, 1999) Vol. 274, No. 42, pp. 29677-29682.

ISSN: 0021-9258.

DOCUMENT TYPE: Article

LANGUAGE: English

SUMMARY LANGUAGE: English

L6 ANSWER 3 OF 4 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI **p53**-mediated transcription of the Werner Helicase gene is modulated by Rb

AB The regulation of Werner's syndrome gene (WRN) expression was studied by characterizing the **cis**-regulatory elements in the promoter region and the **p53**.

trans-activating factors that bind to them. First, we defined the transcription initiation sites and the sequence of the 5' upstream region (2.8 kb) of WRN that contains a number of **cis**-regulatory elements, including 7 Sp1, 9 retinoblastoma control element (RCE), and 14 AP2 motifs. A region consisting of nucleotides -67 to +160 was identified as the principal promoter of WRN by reporter gene assays in HeLa cells, using a series of WRN promoter-luciferase reporter (WRN-Luc) plasmids that contained the 5' - truncated or mutated WRN upstream regions. In particular, two Sp1 elements proximal to the transcription initiation site are indispensable for WRN promoter activity and bind specifically to Sp1 proteins. The RCE enhances WRN promoter activity. Coexpression of the WRN-Luc plasmids with various dosages of plasmids expressing Rb or **p53** in Saos2 cells lacking active Rb and **p53** proteins showed that the introduced Rb upregulates WRN promoter activity a maximum of 2.5-fold, while **p53** downregulates it a maximum of 7-fold, both dose dependently. Consistently, the overexpressed Rb and **p53** proteins also affected the endogenous WRN mRNA levels in Saos2 cells, resulting in an increase with Rb and a decrease with **p53**. These findings suggest that WRN expression, like that of other housekeeping genes, is directed mainly by the Sp1 transcriptional control system but is also further modulated by transcription factors, including Rb and **p53**, that are implicated in the **cell cycle**, cell senescence, and genomic instability.

ACCESSION NUMBER: 1998370638 EMBASE

TITLE: Sp1-mediated transcription of the Werner Helicase gene is

AUTHOR: Yamabe Y.; Shimamoto A.; Goto M.; Yokota J.; Sugawara M.; Furutachi Y.

L6 ANSWER 1 OF 4 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.DUPLICATE 1
 TI Repression of CDK1 and other genes with CDE and CHR promoter elements
 AB during DNA damage-induced G2/M arrest in human cells.
 Entry into mitosis is controlled by the cyclin-dependent kinase CDK1 and
 can be delayed in response to DNA damage. In some systems, such G2/M
 arrest has been shown to reflect the stabilization of inhibitory
 phosphorylation sites on CDK1. In human cells, full G2 arrest appears to
 involve additional mechanisms. We describe here the prolonged (>6 day)
 downregulation of CDK1 protein and mRNA levels following DNA damage in
 human cells. This silencing of gene expression is observed in primary
 human fibroblasts and in two cell lines with functional **p53** but
 not in HeLa cells, where **p53** is inactive. Silencing is
 accompanied by the accumulation of cells in G2, when CDK1 expression is
 normally maximal. The response is impaired by mutations in **cis-**
acting elements (CDE and CHR) in the CDK1 promoter,
 indicating that silencing occurs at the transcriptional level. These
 elements have previously been implicated in the repression of
 transcription during G1 that is normally lifted as cells progress into S
 and G2. Interestingly, we find that other genes, including those for
 CDC25C, cyclin A2, cyclin B1, CENP-A, and topoisomerase II.alpha., that
 are normally expressed preferentially in G2 and whose promoter regions
 include putative CDE and CHR elements are also downregulated in response
 to DNA damage. These data, together with those of other groups, support
 the existence of a **p53**-dependent, DNA damage-activated pathway
 leading to CHR- and CDE-mediated transcriptional repression of various
 G2-specific genes. This pathway may be required for sustained periods of
 G2 arrest following DNA damage.
 ACCESSION NUMBER: 2000101579 EMBASE
 TITLE: Repression of CDK1 and other genes with CDE and CHR
 promoter elements during DNA damage-induced G2/M arrest in
 human cells.
 AUTHOR: Badie C.; Itzhaki J.E.; Sullivan M.J.; Carpenter A.J.;
 Porter A.C.G.
 CORPORATE SOURCE: A.C.G. Porter, MRC Clinical Sciences Centre, Imperial
 College School of Medicine, Hammersmith Hospital, London
 W12 0NN, United Kingdom. andy.porter@csc.mrc.ac.uk
 SOURCE: Molecular and Cellular Biology, (2000) 20/7 (2358-2366).
 Refs: 65
 ISSN: 0270-7306 CODEN: MCEBD4
 COUNTRY: United States
 DOCUMENT TYPE: Journal; Article
 FILE SEGMENT: 029
 LANGUAGE: English
 SUMMARY LANGUAGE: English

L6 ANSWER 2 OF 4 BIOSIS COPYRIGHT 2001 BIOSIS
 TI **p53** negatively regulates cdc2 transcription via the
 AB The **p53** tumor suppressor protein regulates the transcription of
 regulatory genes involved in **cell cycle** arrest and
 apoptosis. We have reported previously that inducible expression of the
p53 gene leads to the **cell cycle** arrest both
 at G1 and G2/M in association with induction of p21 and reduction of
 mitotic cyclins (cyclin A and B) and cdc2 mRNA. In this study, we
 investigated the mechanism by which **p53** regulates transcription
 of the cdc2 gene. Transient transfection analysis showed that wild type
p53 represses whereas various dominant negative mutants of
p53 increase cdc2 transcription. The cdc2 promoter activity is not
 repressed in cells transfected with a transactivation mutant, p5322/23. An
 adenovirus oncoprotein, E1B-55K inhibits the **p53**-mediated

(FILE 'HOME' ENTERED AT 09:05:12 ON 05 DEC 2001)

FILE 'BIOSIS', EMBASE, MEDLINE' ENTERED AT 09:05:39 ON 05 DEC 2001
L1 143956 S CELL CYCLE
L2 71863 S P53 OR CMYC
L3 11807 S L1 AND L2
L4 0 S L3 AND (INTERNAL RIBOSOME ENTRY SEQUENCE? OR IRES)
L5 9 S L3 AND (CIS-ACTING ELEMENT?)
L6 4 DUP REM L5 (5 DUPLICATES REMOVED)

AB

block of induced erythroid differentiation. In this paper we report the effects of aromatic poly amides (TAPP-H, TAPP-B and DAPP-H) and their bromo-derivatives (TAPP-Br, TAPP-Br and DAPP-Br) on cell growth and erythroid differentiation of murine Friend erythroleukemic cells. These compounds are strong inhibitors of serine **proteases**. The results obtained give evidence (a) for inhibition of cell proliferation by all the aromatic poly-amides studied; (b) for stronger antiproliferative activity of the Br-derivatives; (c) for a lack of inhibition of globin mRNA accumulation and hemoglobin synthesis in uninduced cells as well as in Friend cells induced to erythroid differentiation by hexamethylenesuccinamide. These results suggest that this class of antiproliferative compounds exhibits differential effects on **cell-cycle** and differentiation specific genes. Thus, these aromatic poly-amides could be used in experimental therapy without interfering with differentiative processes also when combined with differentiating agents.

ACCESSION NUMBER: 89203563 EMBASE

DOCUMENT NUMBER: 1989203563

TITLE:

Friend erythroleukemic cells treated with aromatic polyamides exhibiting antiprotease activity: Inhibition of cell growth is not associated with block of induced erythroid differentiation.

AUTHOR:

Nastruzzi C.; Feriotto G.; Barbieri R.; Gambati R. Dipartimento di Scienze Farmaceutiche, Università di Ferrara, I-44100 Ferrara, Italy

SOURCE:

Cancer Journal, (1989) 2/10 (335-340). ISSN: 0765-7846 CODEN: CANJFI

COUNTRY:

France

DOCUMENT TYPE:

Journal

FILE SEGMENT:

016

Cancer

025

Hematology

029

Clinical Biochemistry

037

Drug Literature Index

English

French; English

LANGUAGE:

SUMMARY LANGUAGE:

L15 ANSWER 19 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
 TI Friend erythroleukemic cells treated with aromatic polyamides exhibiting
 antiprotease activity: Inhibition of cell growth is not associated with

LANGUAGE: English
 SUMMARY LANGUAGE: English; French
 FILE SEGMENT: 016
 DOCUMENT TYPE: 005
 COUNTRY: France
 SOURCE: Cancer Journal, (1993) 6/6 (322-329).
 ISSN: 0765-7846 CODEN: CANJFI
 CORPORATE SOURCE: Department of Pathology, Université Laval, L'Hotel-Dieu, 11
 AUTHOR: Tetu B.; Brissson J.; Roberge N.
 TITLE: Prognostic significance of tumor DNA content and S-phase
 DOCUMENT NUMBER: 1994028321
 ACCESSION NUMBER: 94028321 EMBASE

addition of adjuvant therapy.
 the prognostic significance of DNA ploidy may be influenced by the
 better prognosis in node-positive breast carcinoma. It also suggests that
 This study re-emphasizes the association of diploidy and low SPF with a
 nuclear grade (p = 0.066) and a younger age (p = 0.069). Conclusions -
 associated with larger tumor size (p = 0.001) and more weakly with high
 expression (p = 0.001) heat-shock protein-27 staining (p =
 absence of progesterone receptors (p = 0.012), Her2/neu oncoprotein
 older age at diagnosis (p = 0.009), poor nuclear grade (p > 0.0001),
 independent of other prognostic factors. Aneuploidy was associated with an
 to adjuvant chemotherapy or hormone therapy (p = 0.0012) and was
 SPF was however only significant in the subgroup of patients not submitted
 advantage of diploidy with low SPF over aneuploidy or diploidy with high
 to aneuploidy or diploidy with high SPF (p = 0.0044). The prognostic
 Diploidy with low SPF was a good marker of better prognosis when compared
 the 761 cases, 320 (42%) were diploid and 441 (58%) were aneuploid.
 follow up ranged from 2.5 to 9.5 years (average: 6.5 years). Results - Of
 others were submitted to adjuvant chemotherapy or hormone therapy. The
 of these patients, 199 (26.1%) received no adjuvant therapy while the
 paraffin embedded material from 761 tumors resected between 1980 and 1986.
 phase fraction (SPF) were evaluated by flow-cytometry on formalin-fixed
 not to adjuvant therapy. Patients and Methods - Tumor DNA content and S
 DNA content in patients with node-positive breast carcinoma submitted or
 AB Purpose - This study was aimed at evaluating the prognostic value of tumor
 cytometry in node-positive breast carcinomas.
 TI Prognostic significance of tumor DNA content and S-phase fraction by flow

L15 ANSWER 18 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
 LANGUAGE: English
 SUMMARY LANGUAGE: English
 FILE SEGMENT: 025
 DOCUMENT TYPE: 022
 COUNTRY: France
 SOURCE: Cancer Journal, (1995) 8/4 (195-200).
 ISSN: 0765-7846 CODEN: CANJFI
 CORPORATE SOURCE: Oncology Research Program, Toronto General Hospital, 67
 AUTHOR: Guin B.; Padua R.A.
 TITLE: progression of leukaemia?

AMERICAN, (1997 Jan-Feb) 3 (1) 13-20.

Journal code: CR8; 9513568. ISSN: 1081-4442.

United States

Journal; Article; (JOURNAL ARTICLE)

English

Priority Journals

199703

Entered STN: 19970407

Last Updated on STN: 19980206

Entered Medline: 19970327

AB

To summarize the majority of p53 mutations in leukemias are point

leukemias?

TI Do p53 mutations have a role in the initiation and progression of

mutations mainly in conserved regions IV and V, although CML patients also

exhibit a notable portion of p53 gene rearrangements as seen more

frequently in solid tumors. A significant proportion of the point

mutations of the p53 gene occur at CpG dinucleotides, sites of DNA

methylation which are believed to be preferentially mutated due to

methylation-induced deamination of 5'-methylcytosine, perhaps further

explaining the localization of p53 mutations to more specific sites within

the functional domains of the gene in many cancers, including leukemia.

Arginine also appears to be a preferential site of mutations (at codon

175, 213, 248, 273 and 282). Arginine is coded for by CGN and hence a site

for methylation, explaining part of its vulnerability to mutations.

Disregarding these CpG preferential sites of mutation, the remaining

mutations appear to be random in occurrence with a slight preference to G

or C basepair mutations in MDS and lymphoma patients which may reflect the

etiology of these diseases. DNA lesions occurring in p53 in hematological

malignancies and solid tumors appear to follow a similar pattern of

localization to the central DNA binding region of the p53 gene. Overall,

the frequency of p53 mutations appears to increase with disease

progression in all neoplasms. A higher rate of p53 mutations are found in

AML patients than in the preceding MDS and in CML p53 mutations are more

common in the blast crisis phase than in the preceding chronic phase or

intermediate accelerated phase. This has also been found to be the case in

solid tumors with an increasing frequency of p53 mutations with

progression and/or metastatic propensity. The rate of p53 mutations has

also been found to be increased in cell lines compared to their parent

tumours, perhaps indicative of the growth advantage provided by such

mutations and of the increased genomic instability imposed on tumours by

p53 mutations. The p53 **cell cycle** checkpoint control

for genetic integrity when mutated appears to allow molecular evolution

with increased rapidly and cell division at a increased rate in the

absence of the p53 G1/S stop signal. Mutations of the p53 gene are also

increased in therapy-related leukemias compared to de-novo leukemias

reiterating the DNA-damaging effects of many of the cancer treatment

regimes. The frequency of p53 mutations in de-novo leukemias are low

compared to most other malignancies. The exception to this appears to be

CML, where mutations are reported to occur in as many as 30% of patients

at blast crisis, some of the more severe forms of ALL and established

leukemic cell lines. It is possible that another mechanism of p53

dysfunction, such as elevated MDM-2 expression may be the cause of the

aberration of normal p53 function in leukemias. Another mechanism of cell

deregulation unrelated to the p53 pathway may otherwise be involved in the

development of hematological neoplasias. It is surprising that one gene

has had such a central role in so many critical processes and hence in the

development of such a diverse range of cancers.

ACCESSION NUMBER: 95261853 EMBASE

DOCUMENT NUMBER: 1995261853

TITLE:

Do p53 mutations have a role in the initiation and

115 ANSWER 15 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
 TI Tumor-related prognostic factors for breast cancer.
 AB Interest in prognostic factors for breast cancer has been stimulated by the success of systemic adjuvant therapy for early-stage operable disease. Patients destined for recurrence can be selected for systemic adjuvant therapy, while patients not likely to recur can be spared the morbidity of unnecessary treatment. The number of tumor-related features available for prognosis has grown impressively in recent years. The purpose of this article is to review tumor-related biologic factors and relate them to prognosis and treatment objectives.

ACCESSION NUMBER: 97034636 EMBASE
 DOCUMENT NUMBER: 1997034636
 TITLE: Tumor-related prognostic factors for breast cancer.
 AUTHOR: Donegan W.L.
 CORPORATE SOURCE: Dr. W.L. Donegan, Department of Surgery, Sinai Samaritan Medical Center, Milwaukee, WI, United States
 SOURCE: *Ca-A Cancer Journal for Clinicians*, (1997) 47/1 (28-51).
 Refs: 174
 ISSN: 0007-9235 CODEN: CAMCAM
 COUNTRY: United States
 DOCUMENT TYPE: Journal; General Review
 FILE SEGMENT: 005 General Pathology and Pathological Anatomy
 010 Obstetrics and Gynecology
 016 Cancer
 017 Public Health, Social Medicine and Epidemiology
 LANGUAGE: English
 SUMMARY LANGUAGE: English

115 ANSWER 16 OF 19 MEDLINE
 TI Inhibition of Raf-1 **protein** kinase by antisense phosphorothioate oligodeoxynucleotide is associated with sensitization of human laryngeal squamous carcinoma cells to gamma radiation.
 AB BACKGROUND: Previous studies suggest a central role for Raf-1 **protein** kinase in mitogenic and radiation damage responsive signaling pathways. PURPOSE: Here we report that the expression and enzymatic activity of Raf-1 **protein** are inhibited in SQ-20B cells exposed to rat antisense oligodeoxynucleotide (As-ODNs) directed against the translation initiation site of human c-raf-1 cDNA. In contrast, treatment of SQ-20B cells with an equimolar concentration of rat sense oligodeoxy-ribonucleotide (S-ODNs) had no effect on the expression and activity of Raf-1. RESULTS: We have observed radiosensitization of rat As-ODNs-treated SQ-20B cells. The dose modifying factor of As-ODNs treatment was approximately 1.4. CONCLUSIONS: These studies demonstrate that rat As-ODNs is a DNA sequence-specific radiosensitizer that may have potential for use in the radiation therapy of cancers.

ACCESSION NUMBER: 97205637 MEDLINE
 DOCUMENT NUMBER: 97205637 Pubmed ID: 9072302
 TITLE: Inhibition of Raf-1 **protein** kinase by antisense phosphorothioate oligodeoxynucleotide is associated with sensitization of human laryngeal squamous carcinoma cells to gamma radiation.
 AUTHOR: Soldatenkov V A; Dritschilo A; Wang F H; Olah Z; Anderson W B; Kasid U
 CORPORATE SOURCE: Department of Radiation Medicine, Lombardi Cancer Center, Georgetown University Medical Center, Washington, D.C.
 CONTRACT NUMBER: CA46641 (NCI)
 CA58984 (NCI)
 CA65012 (NCI)
 +
 SOURCE: CANCER JOURNAL FROM SCIENTIFIC

TI Altered regulation of specific **cell-cycle** genes as a marker for transformed and fetal cells.

ACCESSION NUMBER: 97278885 EMBASE
DOCUMENT NUMBER: 1997278885

TITLE: Altered regulation of specific **cell-cycle** genes as a marker for transformed and fetal cells.

AUTHOR: Hengstschlager M.; Soucek T.; Fusch O.; Braun K.; Bernaschek G.; Hengstschlager-Ottner E.
M. Hengstschlager, Obstetrics and Gynecology, University of Vienna, Dept. of Prenatal Diagnosis/Ther., Wahringgasse 18-20, A-1090 Vienna, Austria

SOURCE: **Cancer Journal**, (1997) 10/4 (207-210).
Refs: 20
ISSN: 0765-7846 CODEN: CANJFI
France
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 016 Cancer
LANGUAGE: English

L15 ANSWER 14 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI Cell division cycle alterations in human malignancies.
AB This paper briefly reviews studies on the genes that control the cell division process in human carcinogenesis. It describes the key biochemical events of the division cycle, their regulation, and genetic alterations that occur in human malignancies. The process of cell division is the result of a cycling and time-regulated activation of specific enzymes that phosphorylate (and thus control) **proteins** required for progression toward cell mitosis. These enzymes, named cyclin-dependent kinases (CDKs), are activated by the binding of **protein** co-factors (i.e. cyclins) and are inhibited by specific **proteins** (CDK-inhibitors or CDIs). A large body of evidence demonstrates that several genes involved in **cell cycle** control are altered in human tumors. These include, in addition to the two archetypal tumor suppressor genes RB and p53, a CDK gene (CDK4), a cyclin gene (CCDN1) and a CDI gene (p16INK4A). In particular, p16INK4A gene is inactivated in many different malignancies with sometimes a greater than 80% incidence. These observations, along with studies carried out in animal models, suggest that almost all human cancers show alterations of the molecular mechanisms that control the division cycle. Such alterations might allow somatic cells to escape physiological cellular differentiation and senescence, thus favouring the development of further genetic aberrations and the progression towards an invasive malignant phenotype. In conclusion, we shall envisage possible scenarios for future clinical research (including potential diagnostic developments and new therapeutic strategies) that are related to the above biochemical investigations.

ACCESSION NUMBER: 97243439 EMBASE
DOCUMENT NUMBER: 1997243439

TITLE: Cell division cycle alterations in human malignancies.
AUTHOR: Della Ragione F.; Borriello A.; Giordani L.; Tolascio A. F. Della Ragione, Inst. of Biochem. of Macromolecules, Medical School, Second University of Naples, Via Costantinopoli 16, 80138 Napoli, Italy

SOURCE: **Cancer Journal**, (1997) 10/3 (151-156).
Refs: 20
ISSN: 0765-7846 CODEN: CANJFI
France
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 016 Cancer
LANGUAGE: English

DOCUMENT TYPE: Journal; General Review
 FILE SEGMENT: 005 General Pathology and Pathological Anatomy
 LANGUAGE: English
 SUMMARY LANGUAGE: English
 L15 ANSWER 10 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
 TI p53 Expression, proliferative activity and prognosis in cancer.
 ACCESSION NUMBER: 1998376320 EMBASE
 TITLE: p53 Expression, proliferative activity and prognosis in cancer.
 AUTHOR: Pich A.
 CORPORATE SOURCE: A. Pich, Dept. of Biomed. Sci./Human Oncology, University of Turin, Via Santena 7, 10126 Torino, Italy.
 SOURCE: *Cancer Journal*, (1998) 11/5 (223-228).
 Refs: 28
 ISSN: 0765-7846 CODEN: CANJFI
 FRANCE
 DOCUMENT TYPE: Journal; General Review
 FILE SEGMENT: 005 General Pathology and Pathological Anatomy
 LANGUAGE: English
 L15 ANSWER 11 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
 TI Dietary restriction, amino acid availability and cancer.
 ACCESSION NUMBER: 1998299937 EMBASE
 TITLE: Dietary restriction, amino acid availability and cancer.
 AUTHOR: Wheatley D.N.
 CORPORATE SOURCE: D.N. Wheatley, Cell Pathology Unit, University Medical School, Foresterhill, Aberdeen AB25 2ZD, United Kingdom.
 SOURCE: *Cancer Journal*, (1998) 11/4 (183-189).
 Refs: 80
 ISSN: 0765-7846 CODEN: CANJFI
 FRANCE
 DOCUMENT TYPE: Journal; General Review
 FILE SEGMENT: 016 Cancer
 LANGUAGE: English
 L15 ANSWER 12 OF 19 BIOSIS COPYRIGHT 2001 BIOSIS
 DUPLICATE 5
 TI Control of eukaryotic cell cycle progression by
 phosphorylation of cyclin-dependent kinases.
 ACCESSION NUMBER: 1998:299160 BIOSIS
 DOCUMENT NUMBER: PREV199800299160
 TITLE: Control of eukaryotic cell cycle progression by phosphorylation of cyclin-dependent kinases.
 AUTHOR(S): Morgan, David O. (1); Fisher, Robert P.; Espinoza, F. Herman; Farrell, Allison; Nourse, Jamison; Chamberlin, Holly; Jin, Pei
 CORPORATE SOURCE: (1) Dep. Physiol, Univ. California, 513 Ramassus Ave., San Francisco, CA 94143 USA
 SOURCE: *Cancer Journal from Scientific American*, (May, 1998) Vol. 4, No. SUPPL. 1, pp. S77-S83.
 ISSN: 1081-4442.
 DOCUMENT TYPE: General Review
 LANGUAGE: English
 L15 ANSWER 13 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.

TITLE: Primary central nervous system tumors: Advances in knowledge and treatment.

AUTHOR: Prados M.D.; Berger M.S.; Wilson C.B.

CORPORATE SOURCE: Prof. M.D. Prados, Department of Neurological Surgery, University of California, San Francisco, CA, United States

SOURCE: *Ca-A Cancer Journal for Clinicians*, (1998) 48/6 (331-360).

Refs: 47

ISSN: 0007-9235 CODEN: CAMCAM

COUNTRY: United States

DOCUMENT TYPE: Journal; Article

FILE SEGMENT: 008 Neurology and Neurosurgery

LANGUAGE: English

SUMMARY LANGUAGE: English

L15 ANSWER 8 OF 19 BIOSIS COPYRIGHT 2001 BIOSIS DUPLICATE 4

TI The role for translational control of the cell cycle.

ACCESSION NUMBER: 1998:498051 BIOSIS

DOCUMENT NUMBER: PREV199800498051

TITLE: The role for translational control of the cell cycle.

AUTHOR(S): Parsels, Leslie A.; Chu, Edward (1)

CORPORATE SOURCE: (1) VA Connecticut Healthcare Syst., Cancer Cent., 111-D, 950 Campbell Ave., West Haven, CT 06516 USA

SOURCE: *Cancer Journal from Scientific American*, (Sept.-Oct., 1998) Vol. 4, No. 5, pp. 287-295.

ISSN: 1081-4442.

DOCUMENT TYPE: General Review

LANGUAGE: English

L15 ANSWER 9 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.

TI The p53 pathway and cancer therapy.

AB The p53 tumor suppressor gene is the major target for genetic alteration or biochemical inactivation in human cancer. Tumors harboring p53 mutations are less likely to respond to chemotherapy and radiotherapy. Recent experiments have begun to uncover the regulation and normal function of p53 in tumor suppression. The p53 pathway can be triggered by cellular stress, including DNA damage, hypoxia and oncogene activation. A number of cellular proteins, including MDM2, ARF, ING1, HIF1.alpha., BCL2, CBF/p300 and various kinases such as DNA-PK and ATM modulate p53 stability and function. The best understood biochemical function of p53 is activation of gene expression, for which the target genes appear to impact on control of the cell division cycle, apoptosis, cellular differentiation, senescence, angiogenesis, DNA repair and genomic stability. These targets include p21WAF1, GADD45, MDM2, KILLER/DR5, bax, p130, 14-3-3, sigma, Tsp1 and IGF-BP3. Novel anti-cancer drug strategies are being developed to replace, restore, bypass or mimic p53 function. In addition, the search is on to identify agents which can selectively kill cancer cells lacking p53-dependent checkpoints.

ACCESSION NUMBER: 1998376321 EMBASE

TITLE: The p53 pathway and cancer therapy.

AUTHOR: El-Deiry W.S.

CORPORATE SOURCE: W.S. El-Deiry, Molec. Oncol./Cell Cycle Regul. Lab., Howard Hughes Medical Institute, Univ. of Pennsylvania School of Med., 415 Curie Blvd, Philadelphia, PA, United States.

SOURCE: *Cancer Journal*, (1998) 11/5 (229-236).

Refs: 41

ISSN: 0765-7846 CODEN: CANJFI

COUNTRY: France

cycle. In normal and cancerous cells, a balance between proapoptotic and antiapoptotic signals exists. **Protein kinase C** (PKC) is a cellular serine/threonine kinase with a central role in the mediation of mitogenic signals as well as the regulation of anti-apoptotic signals. Inhibition of PKC by a novel group of chemical agents (PKC inhibitors) can induce apoptosis in some malignant cell lines, act as differentiating agents, and enhance the effect of cytotoxic chemotherapy. Other kinase inhibitors are potent inhibitors of kinases involved in the control of **cell cycle** progression (cyclin-dependent kinases (cdks)). Cdk inhibitors are able to induce **cell cycle** arrest in neoplastic cells and also act as enhancers of chemotherapy-induced apoptosis. The catalytic domain of different classes of kinases (PKC and cdk) share considerable homology. As a result, many kinase inhibitors that act by blocking the catalytic site are not highly specific and may act as inhibitors of both PKC and cdk. Preliminary studies point to potential applications for some of these PKC/cdk inhibitors, and current clinical trials are exploring the role these agents might have in cancer therapy. In this article, we discuss the rationale for the development of this novel class of agents and highlight those drugs, which have shown promise in clinical testing.

ACCESSION NUMBER: 2001:24657 BIOSIS
DOCUMENT NUMBER: PREV200100024657
TITLE: Cyclin-dependent kinase and **protein kinase C** inhibitors: A novel class of antineoplastic agents in clinical development.

AUTHOR(S): Kaubisch, Andreas; Schwartz, Gary K. (1)
CORPORATE SOURCE: (1) Division of Gastrointestinal Oncology, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY, 10021 USA
SOURCE: **Cancer Journal**, (July-August, 2000) Vol. 6, No. 4, pp. 192-212, print.
ISSN: 1528-9117.

DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English

L15 ANSWER 6 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI Cancer Journal: Editorial.
ACCESSION NUMBER: 2000374876 EMBASE

TITLE: Cancer Journal: Editorial.
AUTHOR: Devita V.T. Jr.; Hellman S.; Rosenberg S.A.
SOURCE: **Cancer Journal**, (2000) 6/4 (191).
ISSN: 1528-9117 CODEN: CAJOCB

COUNTRY: United States
DOCUMENT TYPE: Journal; Editorial
FILE SEGMENT: 016
Cancer
Pharmacology
030
037 Drug Literature Index
LANGUAGE: English

L15 ANSWER 7 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI Primary central nervous system tumors: Advances in knowledge and treatment.
AB The ability to diagnose, monitor, and treat CNS tumors has been improved by new imaging techniques such as positron emission tomography (PET) scanning and functional MR imaging, stereotactic surgery, delivery of radiotherapy with brachytherapy and radiosurgery, and novel methods for delivering chemotherapy. These innovations combined with the new information about tumor pathogenesis and behavior revealed by molecular research give hope that more specific treatments for malignant CNS tumors will be developed in the future.
ACCESSION NUMBER: 1998393029 EMBASE

reverse transcriptase polymerase chain reaction analysis revealed a markedly distinct molecular expression profile in the CL lines: overexpression of basic fibroblast growth factor, interleukin-6, interleukin-8, vascular endothelial growth factor, transforming growth factor-beta, epidermal growth factor receptor, caveolin, and bcl-2 messenger RNAs and marked down-regulation of E-cadherin, p-53, and pentaerythritol tetranitrate. CONCLUSIONS Early administration of hormonal therapy after failure of first line treatment is associated with a profound clonal selection of aggressive AI variants, such as CL-1 and CL-2 lines. These tumor lines, with their parental counterparts, can serve as valuable tools for studying the cellular and molecular mechanisms of CAP progression and metastasis under hormonal therapy. CL-1 and CL-2 offer a unique and reproducible model for the evaluation of drug sensitivity and for other therapeutic modalities for advanced prostate cancer.

ACCESSION NUMBER: 2001:26304 BIOSIS
DOCUMENT NUMBER: PREV200100026304
TITLE: Androgen deprivation induces selective outgrowth of aggressive hormone-refractory prostate cancer clones expressing distinct cellular and molecular properties not present in parental androgen-dependent cancer cells.

AUTHOR(S): Tso, Cho-Lea; McBride, William H.; Sun, Jirong; Patel, Belur; Tsui, Ke-Hung; Pak, Sun H.; Gittlitz, Barbara; Calliaw, Randy; van Ophoven, Arndt; Wu, Lily; deKernion, Jean; Beldegryn, Arie (1)
CORPORATE SOURCE: (1) Department of Urology, UCLA School of Medicine, CHS 66-118, Los Angeles, CA, 90095-1738 USA
SOURCE: *Cancer Journal*, (July August, 2000) Vol. 6, No. 4, pp. 220-233, print.
ISSN: 1528-9117.
DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English

L15 ANSWER 4 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI Thymidylate synthase and the **cell cycle**: What should we believe?
ACCESSION NUMBER: 2000374879 EMBASE
TITLE: Thymidylate synthase and the **cell cycle**
AUTHOR: Dolnick B.J.
CORPORATE SOURCE: Dr. B.J. Dolnick, Dept. of Experimental Therapeutics, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263-0001, United States
SOURCE: *Cancer Journal*, (2000) 6/4 (215-216).
Refs: 6
ISSN: 1528-9117 CODEN: CAJOCB
COUNTRY: United States
DOCUMENT TYPE: Journal; Note
FILE SEGMENT: 016 Cancer
030 Pharmacology
037 Drug Literature Index
LANGUAGE: English

L15 ANSWER 5 OF 19 BIOSIS COPYRIGHT 2001 BIOSIS
TI Cyclin-dependent kinase and **protein kinase C** inhibitors: A novel class of antineoplastic agents in clinical development.
AB Malignant cells have acquired adaptations, which give them a growth and survival advantage over normal cells. One effect of many of these adaptations is that many cancerous cells are less likely to undergo programmed cell death (apoptosis) and, moreover, are resistant to chemotherapy-induced apoptosis. Other features of neoplastic cells are the loss of regulated or orderly progression through the **cell**

arrest point for both cell lines was before the point at which Rb phosphorylation takes place, yet the confluent H630RI cells had threefold higher p21 than confluent H630WT cells. DISCUSSION These data suggest that the 5-FU-resistant H630RI cell lines arrest at a later point in G0/G1 and have a potentially greater capacity for proliferation.

ACCESSION NUMBER: 2001:26305 BIOSIS
DOCUMENT NUMBER: PREV200100026305

TITLE:

cell cycle regulation of the G0/G1 transition in 5-fluorouracil-sensitive and -resistant human colon cancer cell lines.

AUTHOR(S):

McGinn, Cornelius J.; Pestalozzi, Bernhard C.; Drake, James C.; Glennon, M. Clay; Kunugi, Keith; Otterson, G.; Allegra, Carmen J.; Johnston, Patrick G.; Kinsella, Timothy J. (1) (1) Department of Radiation Oncology, University Hospitals of Cleveland, 11100 Euclid Ave, LTR 6068, Cleveland, OH, 44106-6068 USA

SOURCE:

Cancer Journal, (July August, 2000) Vol. 6, No. 4, pp. 234-242. print. ISSN: 1528-9117.

DOCUMENT TYPE:

Article
English
English

LANGUAGE:

SUMMARY LANGUAGE:

English

L15 ANSWER 3 OF 19 BIOSIS COPYRIGHT 2001 BIOSIS
TI Androgen deprivation induces selective outgrowth of aggressive hormone-refractory prostate cancer clones expressing distinct cellular and molecular properties not present in parental androgen-dependent cancer cells.

AB

PURPOSE The mechanism of progression of human prostate cancer (CaP) cells under androgen ablation therapy remains unclear. To study the alternative pathways of CaP cell growth under conditions of androgen deprivation, androgen-independent CaP variants were selected and expanded from an androgen-dependent CaP line via an in vitro androgen deprivation treatment. Cellular and molecular properties of these androgen-independent variants were characterized both in vitro and in vivo and compared with those of their parental androgen-dependent cells. METHODS Androgen deprivation treatment of an androgen-dependent CaP cell line, LNCaP, was carried out by replacing culture medium with RPMI 1640 medium plus 10% charcoal-stripped serum. Cells that survived through the androgen deprivation treatment were harvested and expanded in the androgen-deficient culture medium and were designated CL-1. The CL-1 cells were also recultured in androgen-containing medium and designated CL-2. The growth (cell cycle analysis, 3H-thymidine incorporation assay, growth expansion, and colonization efficiency), expression of CaP-associated markers (semiquantitative reverse transcriptase polymerase chain reaction), interaction with endothelial and bone marrow stromal cells, sensitivity to anticancer agents and radiation (growth inhibition), and tumorigenicity of CL-1 and CL-2 cells were determined and compared with these characteristics in parental LNCaP cells. RESULTS CL-1 and CL-2 cells are fast-growing cells when compared with parental LNCaP cells. They were capable of potentiating the growth of endothelial and bone marrow stromal cells in co-culture experiments and acquired significant resistance to radiation and to anticancer cytotoxic agents (Taxol(R), paclitaxel, vinblastine, and etoposide). In contrast to the poorly tumorigenic parental LNCaP cells, CL-1 and CL-2 lines proved highly tumorigenic, exhibiting invasive and metastatic characteristics in intact and castrated mice or in female mice within a short period of 3 to 4 weeks. No growth supplements (e.g., Matrigel) were needed. When transfected with the green fluorescence **protein** (GFP) gene and transplanted orthotopically in the accessory sex gland, extensive metastatic disease from the primary CaP tumor could be identified in bone, lymph nodes, lung, liver, spleen, kidney, and brain. Semiquantitative

L15 ANSWER 1 OF 19 MEDLINE
 TI Signaling pathways in pancreatic cancer.
 AB Pancreatic cancer is, indisputably, one of the most malignant gastrointestinal tumors. Although the etiology of this disease is unknown, it is clearly linked to alterations in the biologic activities of various signaling molecules. Aberrant signaling activities of growth factors and their receptors, transcription factors, and proteins that control the **cell cycle** have been increasingly implicated in the pathogenesis and dissemination of pancreatic tumors. It is indeed possible that several of these molecules are, in fact, part of a signaling network that has gone awry. This review summarizes some recent advances in an attempt to generate a working model for future investigations.

ACCESSION NUMBER: 2001513706 IN-PROCESS
 DOCUMENT NUMBER: 21445507 PubMed ID: 11561604

TITLE: Signaling pathways in pancreatic cancer.
 AUTHOR: Reddy S A

CORPORATE SOURCE: Department of Gastrointestinal Medical Oncology, University of Texas M.D. Anderson Cancer Center, Houston 77030, USA.
 SOURCE: **CANCER JOURNAL**, (2001 Jul-Aug) 7 (4) 274-86.
 Journal code: DUN; 100931981. ISSN: 1528-9117.
 United States
 Journal; Article; (JOURNAL ARTICLE)

PUB. COUNTRY: English
 FILE SEGMENT: IN-PROCESS; NONINDEXED; Priority Journals
 ENTRY DATE: Entered STN: 20010920
 Last Updated on STN: 20010920

L15 ANSWER 2 OF 19 BIOSIS COPYRIGHT 2001 BIOSIS
 TI **Cell cycle** regulation of the G0/G1 transition in
 AB 5-fluorouracil-sensitive and -resistant human colon cancer cell lines.
 PURPOSE Resistance to 5-fluorouracil (5-FU) has been associated with
 thymidylate synthase (TS) gene amplification and increased TS
 protein levels. Increased TS protein expression has also
 been found to be a significant independent prognostic factor for
 disease-free survival and overall survival in patients treated with
 adjuvant 5-FU-based chemotherapy. In these studies and in our prior
 preclinical studies, TS has been considered a marker of proliferative
 capacity. The purpose of the current study was to further evaluate the
 association between TS levels and **cell cycle**
 regulation, by investigating **cell cycle** kinetics in a
 5-FU-resistant cell line with constitutive overexpression of TS. The
 influence of increased TS levels on **cell cycle**
 progression may provide insight into methods to overcome 5-FU resistance.
 MATERIALS 5-FU-sensitive NCI H630WT and 5-FU-resistant NCI H630R1 (with
 15- to 20-fold higher TS protein levels) were utilized in this
 investigation to determine the influence of constitutive overexpression of
 TS on **cell cycle** kinetics. RESULTS There was no
 apparent influence of increased TS levels on **cell cycle**
 distribution during asynchronous growth, and both cell lines reach plateau
 growth phase in 120 hours, arresting in G0/G1, as determined by flow
 cytometry. In the H630WT cells, this G0/G1 arrest was associated with a
 14- to 17-fold reduction in TS activity and protein levels
 (using the TS-106 monoclonal antibody), whereas in the H630R1 cells, only
 a two- to fivefold reduction was noted. Flow cytometry analysis utilizing
 Ki-67 indicated that there was no evidence of a G0 population in the
 confluent H630R1, whereas 26% +/- 7% of confluent H630WT cells were Ki-67
 negative (G0) and the remainder had low Ki-67 signal intensity. Analysis
 of pRb phosphorylation and p16 and p21 expression suggested that the

FILE 'CANCERLIT' ENTERED AT 09:17:08 ON 05 DEC 2001	L7
2 S (CANCER JOURNAL)/CS	L8
0 S L7 AND 1998<=PY<=1998	L9
0 S CELL CYCLE DEPENDENT TRANSLATION	L10
0 S BIOSIS EMBASE MEDLINE CAPLUS CANCERLIT	
FILE 'BIOSIS, EMBASE, MEDLINE, CAPLUS, CANCERLIT' ENTERED AT 09:20:48 ON	
05 DEC 2001	L11
216006 S CELL CYCLE	L12
5674763 S PROTEIN?	L13
107329 S L11 AND L12	L14
31 S L13 AND CANCER JOURNAL/SO	L15
19 DUP REM L14 (12 DUPLICATES REMOVED)	